

# **Wechselwirkungen zwischen Assemblierungsdomänen von Spinnenseidenproteinen**

## **DISSERTATION**

Zur Erlangung des akademischen Grades

**Doktor der Naturwissenschaften**

(Dr. rer. nat.)

im Promotionsprogramm „Molekulare Biowissenschaften“ der Bayreuther  
Graduiertenschule für Mathematik und Naturwissenschaften (BayNAT) der  
Universität Bayreuth

vorgelegt von

**JOSCHKA BAUER**

aus Nürnberg

Bayreuth 2018





Die vorliegende Arbeit wurde in der Zeit von März 2013 bis September 2017 in Bayreuth am Lehrstuhl Biomaterialien, Fakultät für Ingenieurwissenschaften, unter Betreuung von Herrn Professor Dr. Thomas Scheibel angefertigt.

Vollständiger Abdruck der von der Bayreuther Graduiertenschule für Mathematik und Naturwissenschaften (BayNAT) der Universität Bayreuth genehmigten Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat).

Dissertation eingereicht am: 11.09.2017

Zulassung durch das Leitungsgremium: 28.09.2017

Wissenschaftliches Kolloquium: 20.04.2018

Amtierender Direktor: Prof. Dr. Dirk Schüler

Prüfungsausschuss:

Prof. Dr. Thomas Scheibel (Erstgutachter)

Prof. Dr. Andreas Möglich (Zweitgutachter)

Prof. Dr. Birgitta Wöhr (Vorsitz)

Prof. Dr. Birte Höcker

(Drittgutachter: Prof. Dr. Nico Bruns)



---

**INHALTSVERZEICHNIS**

<b>ZUSAMMENFASSUNG .....</b>	<b>1</b>
<b>SUMMARY .....</b>	<b>4</b>
<b>1. EINLEITUNG .....</b>	<b>6</b>
1.1.    PROTEINE .....	6
1.2.    FIBRILLÄRE PROTEINE .....	6
1.3.    SEIDE .....	11
1.4.    SPINNENSEIDENTYPEN .....	11
1.5.    HIERARCHISCHER AUFBAU DES DRAGLINE FADENS .....	14
1.6.    NATÜRLICHER SPINNPROZESS .....	16
1.7.    SPINNENSEIDENPROTEINE .....	17
1.7.1    NATÜRLICHE SPINNENSEIDENPROTEINE .....	17
1.7.2    REKOMBINANT HERGESTELLTE SPINNENSEIDENPROTEINE .....	18
1.7.3    KERNDOMÄNE .....	19
1.7.4    AMINO-TERMINALE DOMÄNE .....	22
1.7.5    CARBOXY-TERMINALE DOMÄNE .....	31
1.7.6    WECHSELWIRKUNGEN ZWISCHEN REKOMBINANTEN ASSEMBLIERUNGSDOMÄNEN .....	34
<b>2. ZIELSETZUNG .....</b>	<b>37</b>
<b>3. SYNOPSIS .....</b>	<b>38</b>
3.1    ÜBERSICHT ÜBER UNTERSUCHTE ASSEMBLIERUNGSDOMÄNEN VON SPINNENSEIDENPROTEINEN .....	38
3.2    STRUKTURELLE UND FUNKTIONELLE CHARAKTERISIERUNG DER NICHT- REPETITIVEN N-TERMINALEN DOMÄNE NRN .....	41
3.3    WECHSELWIRKUNGEN ZWISCHEN NRN UND KERNDOMÄNE .....	44
3.4    STRUKTURELLE UND FUNKTIONELLE CHARAKTERISIERUNG DER NICHT- REPETITIVEN C-TERMINALEN DOMÄNE NRC .....	48
3.5    WECHSELWIRKUNGEN ZWISCHEN NRC UND KERNDOMÄNE .....	49
3.6    WECHSELWIRKUNGEN ZWISCHEN TERMINALEN DOMÄNEN .....	50
<b>4. LITERATURVERZEICHNIS .....</b>	<b>52</b>
<b>5. PUBLIKATIONSLISTE .....</b>	<b>66</b>

<b>6. DARSTELLUNG DES EIGENANTEILS .....</b>	<b>67</b>
<b>TEILARBEITEN .....</b>	<b>69</b>
<b>DANKSAGUNG.....</b>	<b>146</b>
<b>ERKLÄRUNG .....</b>	<b>148</b>

## Zusammenfassung

Die *Dragline*-Seide der Spinne ist auf Grund der außergewöhnlich hohen Zähigkeit von allen biologischen Fasern die wohl am besten untersuchte. Der *Dragline*-Seidenfaden besteht neben anderen Bestandteilen hauptsächlich aus den *Major ampullate* Spinnenseidenproteinklassen (Spidroinen) MaSp1 und MaSp2. Die Spidroine bestehen aus nicht-repetitiven N- (NRN) und C-terminalen (NRC) Fünf-Helix-Bündeln und einer intrinsisch unstrukturierten und aus repetitiven Modulen aufgebauten Kerndomäne. Die Aminosäuresequenzen der terminalen Domänen sind zwischen den Spidroinen verschiedener Spinnenspezies und Seidenarten hochkonserviert, was deren Wichtigkeit für den Spinnprozess unterstreicht. Das parallel ausgerichtete Dimer von NRC verbindet zwei Spidroine und wird in den meisten MaSp über eine Disulfidbrücke stabilisiert. Dahingegen liegt NRN bei den Lagerungsbedingungen in der Spinndrüse als stabiles Monomer vor. Obwohl beide terminalen Domänen die Löslichkeit der Spidroine während der Lagerung erhöhen, initiieren sie ebenso die Selbstassemblierung der repetitiven Module entlang des Spinntrakts. Hier treten Scherkräfte auf, die hydrophobe Bereiche des NRC-Dimers exponieren und zusammen mit einer ansteigenden Kaliumphosphatkonzentration die intermolekularen Interaktionen zwischen den hydrophoben Modulen der Kerndomäne verstärken. Zum anderen ändert NRN nach der simultanen Absenkung des pH-Wertes und der Natriumchloridkonzentration die Konformation seines Fünf-Helix-Bündels und bildet antiparallele Dimere aus, sodass die über NRC dimerisierten Spidroine zu langen Proteinketten quervernetzt werden.

In dieser Arbeit wurden die Prozesse, die während der Spidroinselbstassemblierung ablaufen, auf molekularer Ebene im Detail analysiert. Die pH- und Salz-induzierte Änderung der Tertiär- und Quartärstruktur (sprich Dimerisierung) von NRN wird über die Neutralisierung von negativ geladenen Carboxylatseitengruppen vermittelt. Um die Rolle der hochkonservierten Asparagin- (D39) und Glutaminsäuren (E76, E81, E114) während der Dimerbildung und Strukturänderung zu identifizieren, wurden einzelne oder multiple Ladungen von NRN<sub>1L.h.</sub> aus *Latrodectus hesperus* MaSp1 durch die Substitution mit den nicht-titrierbaren Analoga Asparagin und Glutamin neutralisiert bzw. durch Austausch mit Arginin umgekehrt. Die resultierenden Varianten wurden rekombinant in dem Wirtsorganismus *Escherichia coli* hergestellt und mittels chromatographischer Trennverfahren gereinigt. Durch Charakterisierung der Varianten mit Mehrwinkel-Lichtstreuungsexperimenten bei Lösungsmittelbedingungen, die während der Lagerung und Assemblierung in der Spinndrüse vorliegen, wurde die Variante E114Q als nahezu pH-

unabhängiges Dimer identifiziert. In Folge dessen steuert die Protonierung der Glutaminsäure E114 die pH-abhängige Dimerbildung. Zudem führte eine Ladungsumkehr in der Dimerisierungsschnittstelle selbst bei Dimer-favorisierenden Bedingungen zu einer signifikanten Stabilisierung der Variante D39R im monomeren Zustand.

Mittels CD-, Fluoreszenz- und NMR-Spektroskopie durchgeführte Strukturanalysen der Varianten von NRN1<sub>L.h.</sub> ergaben, dass die pH-abhängige Strukturänderung von Konformation I zu der eng gepackten Konformation II über die Neutralisierung der kumulierten D39, E76 und E81 gesteuert wird und unabhängig von der Dimerisierung abläuft. Die Struktur einer Dreifachmutante mit neutralisiertem sauren *Cluster* (3\* = D39N-E76Q-E81Q) wurde in Kooperation mit dem Lehrstuhl für Biopolymere (Universität Bayreuth) mittels Kernspinresonanzspektroskopie bestimmt. In der Variante 3\* reduzierte die dauerhafte Neutralisierung der kumulierten negativen Ladungen die elektrostatische Abstoßung zwischen Helix 2 und 3, wodurch das Fünf-Helix-Bündel in der eng-gepackten Konformation II stabilisiert wurde.

Im nächsten Schritt wurde der Zusammenhang zwischen der Dimerisierung von NRN und der Assemblierung der repetitiven Module der Kerndomäne untersucht. Dafür wurden neben NRN1<sub>L.h.</sub> die im monomeren (D39R) und dimeren (3\*\* = E76Q-E81Q-E114Q) Zustand stabilisierten NRN1<sub>L.h.</sub>-Varianten über einen kurzen *Linker* mit 12 Wiederholungen der rekombinanten (AQ)-Module, die nach der Konsensussequenz der Kerndomäne des *Araneus diadematus* Fibroin 3 entworfen wurden, zu dem Konstrukt wtN1L-(AQ)<sub>12</sub> fusioniert. Das pH-abhängige Monomer-Dimer Gleichgewicht des resultierenden wtN1L-(AQ)<sub>12</sub> sowie die Stabilisierung von D39R-(AQ)<sub>12</sub> im monomeren und 3\*\*-(AQ)<sub>12</sub> im dimeren Zustand wurde durch Addition der rekombinanten Kerndomäne nicht beeinflusst. Im Gegensatz dazu unterdrückte die Fusion von NRN1<sub>L.h.</sub> und insbesondere die Addition des im monomeren Zustand stabilisierten D39R die Scherkraft- und KPi-induzierte Assemblierung der repetitiven Module, wohingegen der dimere Zustand die Kinetik der Selbstassemblierung antrieb und die Bildung fibrillärer Aggregate unterstützte.

Während NRN1<sub>L.h.</sub> im sauren pH-Bereich im dimeren Zustand stabilisiert wurde, offenbarte NRC1<sub>L.h.</sub> aus *L. hesperus* MaSp1 eine partielle Entfaltung bei nicht-physiologischen pH-Werten von < 5,0. Dabei begünstigte die Exponierung hydrophober Bereiche die Bildung unstrukturierter Aggregate mit erhöhtem  $\beta$ -Faltblattanteil, allerdings unterschied sich die unspezifische Aggregation von NRC1<sub>L.h.</sub> deutlich von der dynamischen, kontrollierten und reversiblen Assemblierung natürlicher und rekombinanter Spidroine. Während NRC1<sub>L.h.</sub> in dem Fusionskonstrukt NRC1-(AQ)<sub>12</sub> die Scherkraft-induzierte Assemblierung und

Ausrichtung der rekombinanten Kerndomäne im physiologischen pH-Bereich unterstützte, ging die natürliche Funktion als mechano-sensitiver Schalter bei sauren, nicht-physiologischen pH-Werten verloren.

Neben den Strukturen der Termini wurden ebenfalls die Wechselwirkungen zwischen den beiden molekularen Schaltern charakterisiert. Zwischen NRN1<sub>L.h.</sub> und NRC1<sub>L.h.</sub> wurden bei Lösungsmittelbedingungen, die in der Spinndrüse bei der Lagerung und im Spinntrakt bei der Assemblierung der Spidroine vorliegen, keine intermolekularen Interaktionen ausgebildet. Indem der Übergang von löslichen zu assemblierten Spidroinen Ionen-sensitiv über die NRN-Domäne und zudem Scherkraft-induziert über die NRC-Domäne initiiert wird, steuern die terminalen Domänen die Spidroinassemblierung unabhängig voneinander über einen doppelt abgesicherten Kontrollmechanismus.

## Summary

Spider's dragline fiber, best known as one of the toughest biopolymers, is mainly composed of two types of major ampullate spider silk proteins (spidroins), abbreviated as MaSp1 and MaSp2. The spidroins contain non-repetitive N- and C-terminal domains, NRN and NRC, respectively, which are both folded into five-helix-bundles. In contrast, the intrinsically unstructured core domain is made of repetitive modules. The NRN and NRC domains are highly conserved between spider species and silk types, suggesting they play an important role during the spinning process. While NRC permanently connects two spidroins by forming a disulfide-linked, parallel dimer, NRN is monomeric at storage conditions. The terminal domains increase the spidroin's solubility during storage, however, they also trigger the self-assembly. Environmental factors in the spinning duct, which catalyze this process, include shear forces, pH drop and ion swapping. Shear forces induce the exposure of hydrophobic patches of the NRC dimers. Further, the addition of potassium phosphate and removal of sodium chloride strengthen the intermolecular interactions between the hydrophobic core modules. The simultaneous drop of pH and sodium chloride concentration triggers antiparallel dimerization of NRN, physically crosslinking the spidroins into an interconnected network.

In this work, the mechanisms of the spidroin assembly were investigated on the molecular level. The pH-dependent changes of the tertiary and quaternary structure (i.e. dimerization) of NRN are mediated by acidification-induced neutralization of negatively charged carboxylate groups of conserved acidic amino acids. The role of highly conserved aspartic (D39) and glutamic acids (E76, E81, E114) during dimerization and structural conversion was identified by replacing single or multiple acidic residues of NRN1<sub>L.h.</sub> of *Latrodectus hesperus* MaSp1 by their non-titratable analogues asparagine and glutamine or by their basic antagonist arginine. The resulting variants of NRN1<sub>L.h.</sub> were recombinantly produced using *Escherichia coli* as a host system, and subsequently purified using a chromatographic separation strategy. The variants were characterized at solvent conditions that are found during spidroin storage and assembly by performing multi-angle light scattering measurements. Variant E114Q was stabilized as an almost pH-independent dimer, and, accordingly, the pH-sensitive dimer formation is mediated by protonation of glutamic acid E114. Reversing a charge within the dimerization interface significantly stabilized the monomer state even at dimer-favoring solvent conditions.

Structural analysis by CD-, fluorescence and NMR spectroscopy exhibited that the pH-induced structural conversion between conformation I and the tightly packed



conformation II is mediated independently of the dimerization by protonation of the clustered acidic residues D39, E76 and E81. The NMR structure of a triple mutant with neutralized acidic cluster was determined in cooperation with the Department of Biopolymers (University of Bayreuth). Neutralizing the accumulated charges reduced the electrostatic repulsion between helix 2 and 3, which induced the reorganization of the five-helix bundle into the tightly packed conformation II.

In the next step, the relation between dimerization of NRN and assembly of the repetitive core modules into fibrillary assemblies was studied. Wild-type NRN1<sub>L.h.</sub> as well as stable monomers (D39R) and dimers (3\*\* = E76Q-E81Q-E114Q) were fused via a short linker (L) with twelve repeats of a previously established engineered core module (AQ) designed based on the consensus sequence of *Araneus diadematus* fibroin 3. The pH-dependent monomer-dimer equilibrium of the resulting wtN1L-(AQ)<sub>12</sub> as well as stabilization of D39R-(AQ)<sub>12</sub> in the monomeric and 3\*\*-(AQ)<sub>12</sub> in the dimeric state was not affected by addition of the recombinant core domain. However, the shear- and KPi-induced assembly of the repetitive modules was significantly suppressed by fusion to NRN1<sub>L.h.</sub>, and, especially by stabilization of D39R in the monomeric state. In contrast, stabilizing the dimeric state accelerated the assembly kinetics and significantly increased the self-assembly into fibrillary aggregates.

NRN1<sub>L.h.</sub> was stabilized at acidic pH in the dimeric state, whereas the five-helix bundle of NRC1<sub>L.h.</sub> of *L. hesperus* MaSp1 partially unfolded at a probably non-physiological pH of < 5.5. The exposure of hydrophobic patches induced the formation of unstructured aggregates with increased  $\beta$ -sheet content. The aggregation of NRC1<sub>L.h.</sub> could be clearly distinguished from self-assembly of natural and recombinant spidroins, as fibers formed through this process were well-defined and the assembly was reversible. Fusion of NRC1<sub>L.h.</sub> to (AQ)<sub>12</sub> triggered and aligned the shear-induced self-assembly of the recombinant core modules at physiological pH values, however, the natural function of NRC was lost at low pH.

This thesis further investigated the interplay between the two molecular switches NRN and NRC. However, NRN1<sub>L.h.</sub> and NRC1<sub>L.h.</sub> did not interact under solvent conditions mimicking those found during spidroin storage in the sac or assembly in the spinning duct. Since NRN changes its structure in response to the ionic environment and NRC responds to shear stress, this behavior represents an independent double-safety mechanism to prevent premature aggregation and to enable the tight control of spidroin assembly.

## 1. EINLEITUNG

### 1.1. PROTEINE

Lebende Systeme zeichnen sich durch die hochkontrollierte und reproduzierbare Selbstassemblierung ihrer molekularen Strukturen aus. Das fundamentalste und universellste Beispiel der biologischen Selbstassemblierung ist die Faltung von Proteinen in kompakte, drei-dimensionale, globuläre Strukturen,<sup>1</sup> um biologische Funktionen wie beispielweise eine Katalyse oder molekulare Erkennung ausführen zu können.<sup>2</sup> Abhängig von der Primärstruktur bestehen die Proteine vorwiegend aus den Sekundärstrukturelementen  $\alpha$ -Helix,  $\beta$ -Faltblatt, *Random coil* (englisch, unstrukturiert) und deren Mischformen.<sup>3</sup> Die große Vielfalt an hochspezifischen Proteinfaltungen, bei denen funktionelle Gruppen in räumliche Nähe gebracht werden, ermöglichte lebenden Systemen eine erstaunliche Diversität und Selektivität ihrer zugrundeliegenden biologischen Prozesse zu entwickeln.<sup>4</sup>

Die SCOP2 (*Structural Classification of Proteins*) Datenbank<sup>5</sup> teilt die Proteine neben der Proteinkategorie „globulär“ auch in „membranständig“, „intrinsisch ungefaltete“ und „fibrillär“ ein, wobei jede Proteingruppe charakteristische Merkmale aufweist. Membranproteine ermöglichen vorwiegend den Austausch von Molekülen und Informationen über die Zellmembran.<sup>2</sup> Im Gegensatz dazu ist die Funktion von zahlreichen intrinsisch ungefalteten Proteinen noch unbekannt. Während einige unstrukturierte Proteinsegmente nach Bindung an das biologische Zielmolekül eine Faltung einnehmen, sind andere ungefaltete Bereiche als flexibler Linker an der Assemblierung in makromolekulare Strukturen beteiligt.<sup>6</sup> Klassische Faserproteine sind meistens durch repetitive Strukturelemente und eine ausgedehnte, drei-dimensionale Faltung charakterisiert.<sup>3, 7, 8</sup> Die fibrillären Proteine werden auf Grund ihrer Starrheit als nichtkatalytische, makroskopische Strukturen eingesetzt, um Motilität, Elastizität, Stabilität sowie Schutz einer Zelle, einem Gewebe und Organismen zu gewährleisten.<sup>7, 9</sup> Interessanterweise können zunächst globulär gefaltete Proteine wie das Humanalbumin unter moderaten Temperatur-, Salz- oder Druckschwankungen zu faserartigen Proteinen transformieren.<sup>10</sup>

### 1.2. FIBRILLÄRE PROTEINE

Die zunächst ungeordneten, fibrillären Proteine bilden während der Selbstassemblierung untereinander spezifische Interaktionen aus, sodass sich die Proteine in größere Faserstrukturen organisieren.<sup>11</sup> Die fibrillären Proteine unterscheiden sich jedoch

hinsichtlich des zugrundeliegenden Selbstassemblierungsmechanismus sowie der Morphologie und der Verwendung der fibrillären Strukturen (Tabelle 1).

Beispielsweise wird die Selbstassemblierung von Aktin-Untereinheiten über die Bereitstellung freier Filament-Enden initiiert, die durch das Freilegen von bestehenden Filament-Enden, dem Durchtrennen eines Filaments oder der Bildung eines neuen Filament-Nukleus entstehen.<sup>12-14</sup> In der Elongationsphase assoziieren die Untereinheiten an dem schnell wachsenden Ende (*barbed end*, englisch) des Filaments mit einer höheren Bindungsaffinität als an dem langsam wachsenden Ende (*pointed end*, englisch),<sup>15</sup> bis ein stationärer Zustand (*steady state*, englisch) erreicht wird.<sup>12</sup> Die reversible Selbstassemblierung von Aktin-Untereinheiten wird über Monomer- und Filament-bindende Proteine gesteuert, die die Verfügbarkeit von freien Monomeren steuern und die Bildung, Anordnung und Depolymerisation der Filamente regulieren.<sup>14, 15</sup>

Elastische Elastinfasern entstehen *in vivo* durch die Zusammenlagerung von monomeren Tropoelastin (löslicher Präkursor) in kugelförmige Koazervate und deren Ablagerung an vorgeformten Templaten aus Fibrillin-reichen Mikrofibrillen, sodass der Elastinkern von Mikrofibrillen ummantelt wird. Das Elastin enthält alternierende Wiederholungen von hydrophoben und quervernetzenden Domänen, deren Lysin-Seitenketten im koazervierten Zustand durch Lysyloxidasen quervernetzt werden.<sup>16, 17</sup> Die hydrophoben Domänen verleihen dem Elastin die intrinsische Fähigkeit zur Fibrillisierung, sodass auch *in vitro* über spezifische Interaktionen geordnete Fibrillen entstehen.<sup>18, 19</sup>

Die Strukturelemente des Zytoskeletts bestehen vorwiegend aus Intermediärfilamenten, zu deren größten Untergruppe die Keratine zählen. Während der Selbstassemblierung in hierarchisch aufgebaute Keratinfasern steuern die terminalen Proteinbereiche zunächst die Assoziation der Keratin-Untereinheiten des Typs I und II zu *coiled-coil* Heterodimeren, die sich in den nächsten Schritten zu einem symmetrisch aufgebauten Tetramer und weiter zu einem aus 32 Monomeren bestehenden *unit-length filament* (ULF) zusammenlagern.<sup>20-22</sup> Die ULFs assoziieren längsorientiert zu ausgestreckten Fasern, die nach weiteren Reorganisationsprozessen eine kompaktere Form einnehmen und zu einem Netzwerk quervernetzt werden.<sup>22</sup>

Das in nicht-fibrillärer und fibrillärer Form auftretende Kollagen ist eines der in Lebewesen am häufigsten vorkommenden Proteine.<sup>11, 23</sup> Eukaryoten verarbeiten mehr als 20 verschiedene Kollagene, die in mehrere Kollagentypen unterteilt werden, extrazellulär in Bindegewebe, Haut und Knochen.<sup>24</sup> Die Zellen synthetisieren das Kollagen zunächst mit

nicht-helikalen Termini in der Form des löslichen Prokollagens (Präkursor).<sup>23</sup> Die Assemblierung beginnt mit der Anlagerung von drei Kollagenproteinen zu einer charakteristischen Triplehelix. Nach der enzymatischen Abspaltung der amino- und carboxy-terminalen, nicht-helikalen Propeptide im extrazellulären Raum assemblieren die Triplehelices zur Kollagenfaser, die anschließend durch kovalente Quervernetzung weiter stabilisiert wird.<sup>25</sup>

Das Flagellum von Bakterien und Archaeen gleicht einem helikalen Propeller, der es ermöglicht die Zelle anzutreiben.<sup>26</sup> Während der Selbstassemblierung werden die löslichen Flagellin-Moleküle durch den in der cytoplasmatischen Membran verankerten und aus elf Protofilamenten gebildeten, zentralen Kanal zum distalen Ende des Flagellums transportiert. Nach einer strukturellen Änderung assembliert die Flagellin-Untereinheit an den helikalen und längsorientierten Protofilamenten.<sup>27</sup>

Neben der Selbstassemblierung nützlicher Faserproteine kann die ungewollte Selbstassemblierung amyloider Proteine zu zahlreichen, unheilbaren Krankheiten wie z.B. Alzheimer führen.<sup>28, 29</sup> Basierend auf experimentellen und simulierten Studien entstanden zwei Hypothesen über die Selbstassemblierung amyloider Proteine.<sup>30</sup> In dem ersten Szenario ändern die zunächst löslichen, amyloiden Proteine unter leicht denaturierenden Bedingungen partiell die Struktur und nehmen eine unstrukturierte oder alternative Faltung ein. Anschließend lagern sich die Proteine zu einem Nukleus zusammen, der die Struktur der löslichen Proteine ändert und ab einer bestimmten Größe Monomere bindet.<sup>31</sup> In der darauf folgenden Elongationsphase wächst die amyloide Fibrille an beiden Enden.<sup>32, 33</sup> In dem zweiten Modell entstehen aus den amyloiden Proteinen zunächst intermediäre Protofilamente, die zu der Amyloidfibrille assemblieren und an beiden Enden Proteinuntereinheiten binden.<sup>31, 34, 35</sup>

Die zu den Arthropoda (Gliederfüßer) zählenden Klassen Insecta (Insekten), Myriapoda (Tausendfüßer) und Arachnida (Spinnentiere) prozessieren die Faserproteinen zu Seidenfäden, die im exoskeletalen Bereich unterschiedliche Aufgaben übernehmen.<sup>36</sup> Die Seidenproteine falten bei niedrigen Konzentrationen in ihre spezifische Struktur, lagern sich in der Spinndrüse bei höheren Konzentrationen in größere Strukturen mit flüssig-kristallinen Charakter zusammen und assemblieren in fibrilläre Strukturen nachdem sich bestimmte Umgebungsparameter (beispielsweise Wasserentzug) ändern.<sup>37</sup>

**Tabelle 1: Selbstassemblierung von fibrillären Proteinen.** Erweitert und modifiziert nach Scheibel *et al.*<sup>9</sup>

<b>Protein</b>	<b>Aktin</b>	<b>Amyloide Proteine</b>	<b>Elastin</b>
<b>Vorkommen</b>	Intrazellulär in allen eukaryotischen Spezies	Extrazellulär in Organen und Geweben	Extrazellulär in Vertebraten und Insekten
<b>Funktion</b>	Veränderung der Form und Bewegung, Aufnahme von extrazellulärem Material durch Phago- und Endozytose	Ungewollte, pathogene Effekte	Elastisches und widerstandsfähiges Bindegewebe
<b>Fasermorphologie</b>	Polare Struktur, teilweise verzweigt, 7 nm Durchmesser, Länge bis $\mu\text{m}$	Unverzweigt, Cross- $\beta$ Struktur, 10-20 nm Durchmesser	Quervernetztes Netzwerk aus Fasern
<b>Selbstassemblierung</b>	Reversible Addition von monomeren Aktin an einem schnell und langsam wachsenden Ende des Filaments	Nukleus initiiert über eine Strukturänderung der amyloiden Proteine die nicht-reversible Assemblierung zur Fibrille, alternativ Zusammenlagerung von Protofilamenten	Tropoelastin koalesziert in Tropfen, die sich an Mikrofilamenten anlagern und über Lysin-Seitenketten quervernetzt werden
<b>Steuerung der Selbstassemblierung</b>	Monomer- und Filament-bindende Proteine	Initiierung über Nukleus oder Protofilament	Interaktion mit Oberfläche (Rezeptoren) der Bindegewebszellen

<b>Protein</b>	<b>Flagellin</b>	<b>Kollagen</b>	<b>Keratin</b>	<b>Seidenproteine</b>
<b>Vorkommen</b>	Extrazellulär in Prokaryoten und Archaeen	Extrazellulär und exoskeletal in den meisten Eukaryoten	Zytoplasmatisch, extrazellulär und exoskeletal in den meisten Lebewesen	Exoskeletal in Arthropoda
<b>Funktion</b>	Flagellum ermöglicht Zellbeweglichkeit	Strukturelles Gerüst von Gewebe wie z.B. Bindegewebe, Knochen und Haut	Stabilisiert Epithelzellen, Strukturelement in Haaren, Nägeln und Federn	Verwendung der Seide als <i>Life-line</i> (englisch, Rettungsfaden), Netz, Falle und Kokon
<b>Fasermorphologie</b>	Flagellum besteht aus 11 Protofilamenten mit einem Durchmesser von 12-25 nm und eine Länge von bis zu 15 µm, zentraler Kanal	1,4 nm breite und 300 nm lange Triplehelices assemblieren zu Fasern	Nicht-polare Struktur, 10 nm breite und mehrere 100 nm lange Intermediärfilamente	Durchmesser von wenigen hundert nm bis zu 50 µm
<b>Selbstassemblierung</b>	Flagellin wird durch Kanal des Flagellum transportiert und assembliert am distalen Ende des Filaments	Assemblierung und Quervernetzung der Triplehelices	Schrittweise Assoziation des Keratins (Typ I und II) zu Dimeren, Tetrameren, <i>unit-length filaments</i> und Fasern, Reorganisation und Quervernetzung der Fasern	Flüssigkristall-ähnliche Spinnlösung wird in Spinnrüse gelagert, Assemblierung gleicht einer Flüssig-Kristallin-Phasenseparation
<b>Steuerung der Selbstassemblierung</b>	Sekretion ins Flagellum, strukturelle Änderung in der Flagellin-Untereinheit	Enzymatische Abspaltung von terminalen Propeptiden	Interaktionen der terminalen Proteinbereiche	Umgebungsänderungen entlang der Spinnrüse verändern Proteinstruktur

### 1.3. SEIDE

Die zu dem Stamm Arthropoda gehörenden Organismen zeichnen sich durch das Verspinnen löslicher Faserproteine zu unlöslichen Seidenfäden aus.<sup>38</sup> Dabei weisen die homologen Seide-sekretierende Systeme der Insekten und Spinnentiere einige Gemeinsamkeiten auf, wie beispielsweise die Lagerung der Seidenproteine bei hohen Konzentrationen in spezialisierten Drüsen und die Verwendung von Proteinen mit einer hochrepetitiven und Alanin-, Serin- und/oder Glycin-reichen Sequenz. Die einzigartige Primärstruktur der Seidenproteine kodiert für eine bestimmte Mischung aus vorwiegend  $\alpha$ -Helices und  $\beta$ -Faltblättern, die ein makromolekulares Komposit aus amorphen und kristallinen Bereichen entstehen lässt.<sup>39-41</sup> Die semikristalline Beschaffenheit ermöglicht die hohe Zähigkeit des mechanisch stabilen Seidenfadens.<sup>36, 42</sup>

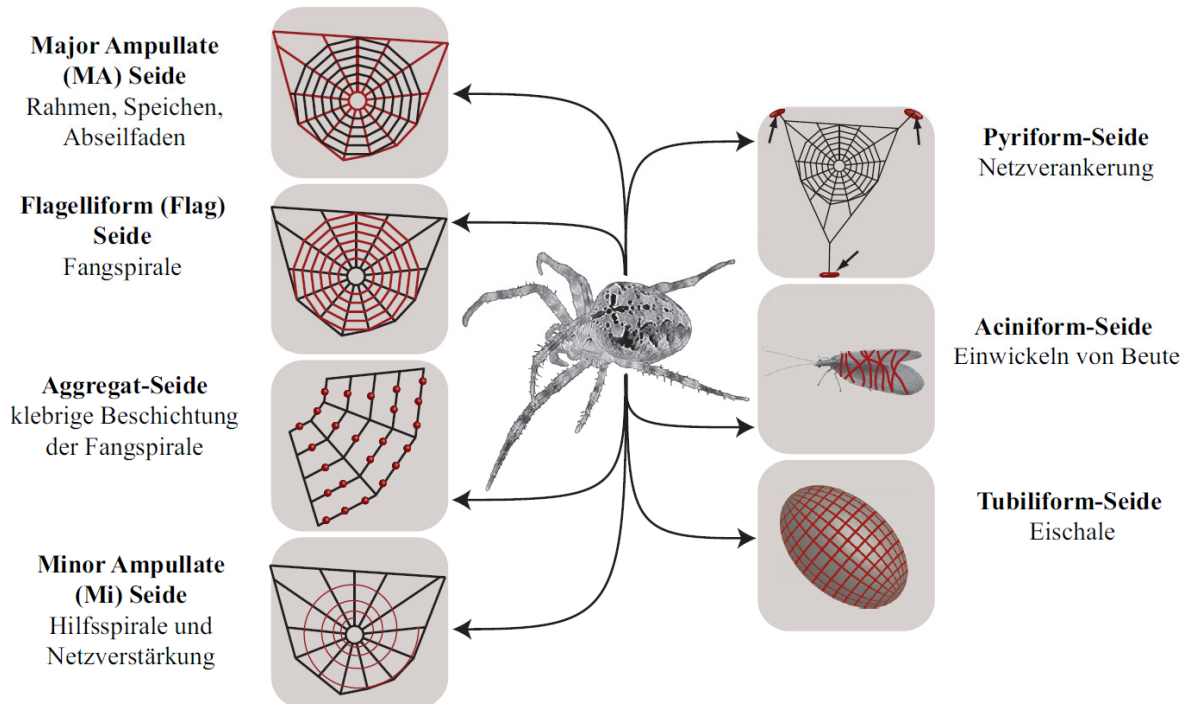
Trotz der strukturellen Ähnlichkeiten haben die Arthropoda im Laufe der Evolution eine große Biodiversität der Seide entwickelt, die den Tieren den variablen Einsatz der Seide in verschiedenen Gebieten ermöglicht.<sup>36</sup> Die Insekten nutzen die Seide typischerweise als Schutznest und zur Herstellung des Kokons.<sup>42</sup> Das wohl bekannteste Beispiel hierfür stellt die Kokonseide von den Seidenspinnern *Bombyx mori* auf Grund der etablierten Züchtung sowie der industriellen Weiterverarbeitung der Seidenfäden dar. Während des fünften Larvenstadiums sekretiert *B. mori* die Faserproteine, die eine leichte Kette Fibroin, eine schwere Kette Fibroin und das Glykoprotein P25 umfassen.<sup>43</sup> Die aus den Fibroinen gebildeten Fasern werden mit klebrigen Glykoproteinen (Sericinen) beschichtet,<sup>44</sup> um die Kohäsion der Seidenfäden im Kokon zu verbessern.<sup>45, 46</sup>

Im Gegensatz zu den Insekten produzieren die Spinnentiere ein breites Spektrum an verschiedenen Seidenfasern, die beispielsweise während des Beutefangs und zum Schutz des Nachwuchses eingesetzt werden.<sup>47</sup>

### 1.4. SPINNENSEIDENTYPEN

Um den spezifischen Anforderungen der jeweiligen Nutzung gerecht zu werden, produzieren Radnetzspinnen wie die Gartenkreuzspinne *Araneus diadematus* neben einem Seidenkleber bis zu sechs verschiedene Seidenfäden (Abbildung 1), die sich in ihrer Zusammensetzung und den daraus resultierenden mechanischen Eigenschaften unterscheiden.<sup>39, 48-51</sup> Obwohl alle Drüsen einen gemeinsamen Ursprung haben,<sup>52, 53</sup> führte die mit der Evolution der letzten 380 Millionen Jahren verbundene Auswahl an spezifischen Funktionalitäten zu einer Biodiversität der Spinnenseide.<sup>54</sup> Die Vielfältigkeit zwischen den Seidentypen lässt sich anhand der divergenten mechanischen Eigenschaften erkennen<sup>55, 56</sup> und beruht auf Unterschieden in der biochemischen Zusammensetzung<sup>57-61</sup>. Die Seidentyp-spezifische

Mischung an Seidenproteinen (Spidroinen) wird selektiv von speziellen Endothelzellen in morphologisch unterschiedlichen Seidendrüsen sekretiert, die namensgebend für die daraus gesponnene Seidenart ist. Die Drüsen können jedoch auch kleinere Mengen an Seidenproteinen produzieren, die vorwiegend von anderen Drüsen hergestellt werden.<sup>58, 62</sup> Im Folgenden wird detailliert auf die spezifische Funktion, die biochemischen Zusammensetzung und die mechanischen Eigenschaften der verschiedenen Seidentypen eingegangen.



**Abbildung 1: Schematische Darstellung der sechs verschiedenartigen Seidenfäden und eines Seidenklebers, die von einer weiblichen Radnetzspinne produziert werden.** Modifiziert nach Eisoldt *et al.*<sup>63</sup> mit freundlicher Genehmigung des Verlages Elsevier.

Das Radnetz besteht aus fünf verschiedenen Seidenarten, die sich entsprechend ihrer Verarbeitung und Funktion im Netz in ihren mechanischen Eigenschaften unterscheiden.<sup>55, 56</sup> Der äußere Rahmen und die Speichen des Netzes bestehen aus der extrem zähen Major-Ampullate (MA) Seide.<sup>54, 64-67</sup> Da die Spinne diese Seide hinter sich herzieht, um sie bei Bedarf als *Life-line* (englisch, Rettungsfaden) zur Flucht vor Angreifern einzusetzen, wird die MA auch *Dragline* (englisch *drag* = ziehen) Seide genannt.<sup>54, 66</sup> Dabei führt das Formgedächtnis der MA Seide zu einzigartigen Torsionsqualitäten, die verhindern, dass sich die Spinne beim Abseilen verdreht und ins Schwingen gerät.<sup>68, 69</sup> Die MA Seide der Gartenkreuzspinne (*Araneus diadematus*) weist eine immens hohe Zähigkeit ( $167 \text{ MJ/m}^3$ ) auf, die aus der außergewöhnlichen Kombination von hoher Zugfestigkeit (1,1 GPa) und moderater Dehnbarkeit (27 %) hervorgeht<sup>55, 70</sup> Diese Zähigkeit wird von den *Dragline*-



Fäden der Species *Gasteracantha cancriformis* ( $177 \text{ mJ/m}^3$ ),<sup>71</sup> *Latrodectus hesperus* ( $181 \text{ mJ/m}^3$ ),<sup>72</sup> *Scytodes* ( $230 \text{ mJ/m}^3$ )<sup>71</sup> und *Caerostris darwini* ( $354 \text{ mJ/m}^3$ )<sup>73</sup> sogar noch übertroffen. Die außergewöhnlichen mechanischen Eigenschaften machen den *Dragline*-Faden von allen biologischen Fasern zu einem der interessantesten und am besten untersuchtesten.<sup>54, 66, 74</sup>

Auf den Rahmen und die Speichen des Radnetzes wird die Flagelliform-Seide als spiralförmiges Fangnetz angebracht. Um die beim Aufprall eines Insektes entstehende kinetische Energie dissipieren zu können, benötigt die Flagelliform-Seide eine Dehnbarkeit von über 200 %. Die Grundlage hierfür bilden die stark amorphen Bereiche des Seidenfadens, die größtenteils aus  $\beta$ -Spiralen (multiple  $\beta$ -Schleifen des Typs II) bestehen und auf der Glycin- und Prolin-reichen Sequenz der Flagelliform-Spidroine basieren.<sup>55, 75, 76</sup>

Um die von der Flagelliform-Seide aufgefangene Beute dauerhaft im Netz zu fixieren, beschichten die cribellaten Spinnen (besitzen kein Cribellum, wie z.B. *Latrodectus hesperus*) das Fangnetz tropfenweise mit der klebrigen und hochviskosen Aggregat-Seide, die eine Mischung aus hygroskopischen Peptiden, Lipiden und phosphorylierten Glykoproteinen mit klebenden Seitenketten enthält. Die cribellaten Spinnen modifizieren das Fangnetz in einem zeitintensiveren Verfahren mit tausenden, zylindrisch geformten und trockenen Nanofasern.<sup>77-81</sup>

Zur weiteren Verstärkung des Netzes wird die Minor-Ampullate (MI) Seide spiralförmig aufgebracht. Obwohl die Major- (MaSp) und Minor-Ampullate Spidroine (MiSp) eine ähnliche Aminosäuresequenz und Sekundärstruktur aufweisen, führt eine ungleiche Nachbehandlung der Seidenfäden zu unterschiedlichen mechanischen Eigenschaften.<sup>82-84</sup> Während sich die MA Seide in wässriger Umgebung um bis zu 50 % der ursprünglichen Länge (Superkontraktion) verkürzt, verändert sich die Länge der MI Seide bei hoher Luftfeuchtigkeit nicht.<sup>85, 86</sup> Da sich die Mikrostrukturen und mechanischen Eigenschaften der beiden Seidenfasern jedoch im trockenen Zustand ähneln, scheint die Assemblierung der MI und MA Spidroine durch eine unterschiedliche Nachstreckung während des Spinnprozesses beeinflusst zu werden.<sup>85</sup>

Während die MA-, MI- und Flagelliform-Seide das Gerüst des Radnetzes bilden, wird die Piriform-Seide zum Verkleben der Seidenarten untereinander und zur Befestigung der Seidenbündel an verschiedenen Fundamenten wie Steinen oder Ästen eingesetzt.<sup>87</sup> Die Piriform-Spidroine weisen durch das Auftreten zweier besonders polarer Blocksequenzen (Serin-, Glutamin- und Alanin-reiches Modul; Block aus Polyprolin) eine deutliche

Sequenzdivergenz zu anderen Seidenarten auf.<sup>88</sup> Der simultane Einsatz zweier Proteinklassen lässt ein Komposit aus faserartig assemblierten Mikrostrukturen und einer Kleber-ähnlichen Lösung entstehen, das im Zusammenspiel mit den langen Fasern der anderen Seidentypen ein kräftiges Netzwerk mit adhäsiven Eigenschaften ausbildet.<sup>88-90</sup>

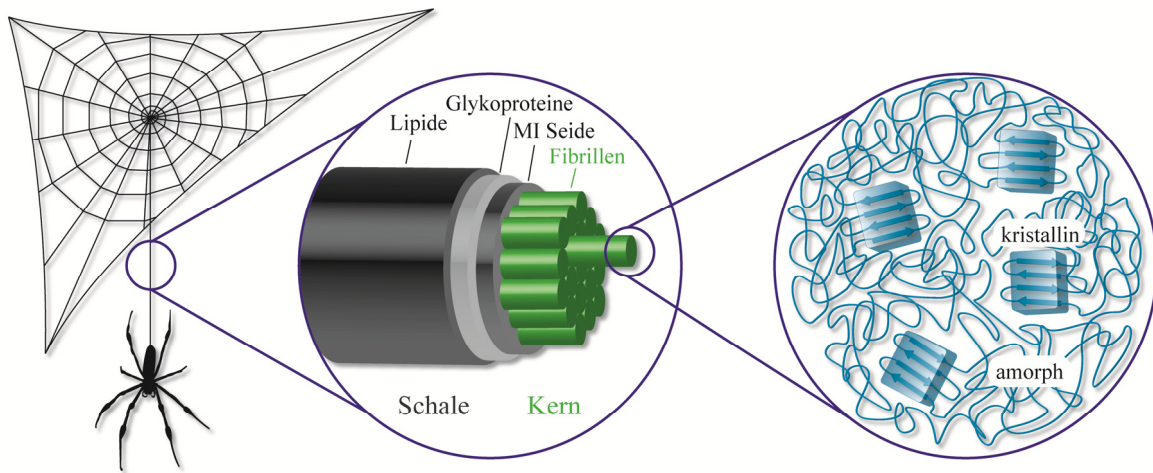
Zwei der sieben Seidenarten der Radnetzspinnen werden während der Fortpflanzung für die Herstellung eines Kokons benötigt. Die äußere Schale des Kokons besteht aus der Tubuliform-Seide und wird auch Cylindriform-Seide genannt. Dieser Seidentyp sorgt für die nötige Stabilität der Struktur und schützt den Nachwuchs vor Angreifern und Parasiten.<sup>58</sup> Verglichen mit den anderen Seidenarten weist die Tubuliform-Seide eine relativ komplexe Aminosäuresequenz mit einem erhöhten Anteil an Serin und Alanin sowie einem reduzierten Glycingehalt auf.<sup>91-93</sup> Zusätzlich enthaltene Cysteine scheinen die Tubuliform-Seidenfasern über kovalent gebundene Disulfidbrücken zu Trimeren quervernetzen.<sup>94, 95</sup> Das Kokoninnere besteht aus der Aciniform-Seide, die darüber hinaus auch bei der Jagd zum Einhüllen und Fixieren der Beute und bei der Fortpflanzung zum Ablegen der Samen eingesetzt wird.<sup>96, 97</sup> Um die mechanischen Anforderungen erfüllen zu können, weist die Aciniform-Seide die höchste Zähigkeit von allen Seidenarten auf.<sup>98, 99</sup>

### 1.5. HIERARCHISCHER AUFBAU DES DRAGLINE FADENS

Die hohe Zugfestigkeit und Elastizität des *Dragline*-Seidenfadens ergibt sich zum einen aus den zugrundeliegenden MaSp1 und MaSp2 und zum anderen aus dem hierarchischen Aufbau.<sup>100, 101</sup> Die komplexe Morphologie des *Dragline*-Fadens umfasst fünf verschiedenen Schichten (Abbildung 2), die sich in ihren Bestandteilen unterscheiden und für verschiedene Eigenschaften des Seidenfadens verantwortlich sind.<sup>100</sup> Vereinfacht kann der Aufbau des Fadens mit einer Kern-Schale Struktur beschrieben werden,<sup>102</sup> bei der die dreiteilige Schale aus einer Lipid-, Glykoprotein- und schwach glykosylierten Proteinschicht besteht. Der zweiteilige Kern enthält entlang der Faserachse ausgerichtete Proteinfibrillen, die hinsichtlich ihrer divergenten Proteinstruktur in eine innere und äußere Region eingeteilt werden.<sup>100, 103</sup>

Die äußerste der fünf Schichten setzt sich ausschließlich aus Lipiden zusammen, die den Seidenfaden gegen äußere Umwelteinflüsse sowie Mikroorganismen schützt und als Trägermaterial für Spezies- und Geschlechts-spezifische Pheromone dient.<sup>104</sup> Auf Grund der lockeren Befestigung an den inneren Schichten werden die mechanischen Eigenschaften des Fadens von der Lipidschicht kaum beeinflusst.<sup>100</sup> Die darunter liegende Schicht aus Glykoproteinen ist deutlich fester an den restlichen Faden gebunden und stellt somit die eigentliche Grenzfläche zur Umgebung dar. Neben dem Schutz vor Mikroorganismen

regulieren die Glykoproteine den Wasserhaushalt und die daraus resultierende Kontraktion des Fadens.<sup>105, 106</sup> Die dritte und innerste Schicht der Schale besteht aus MI-Seidenproteinen und ist neben dem Schutz vor äußeren Einflüssen auch für die Plastizität des Fadens verantwortlich.<sup>100</sup>



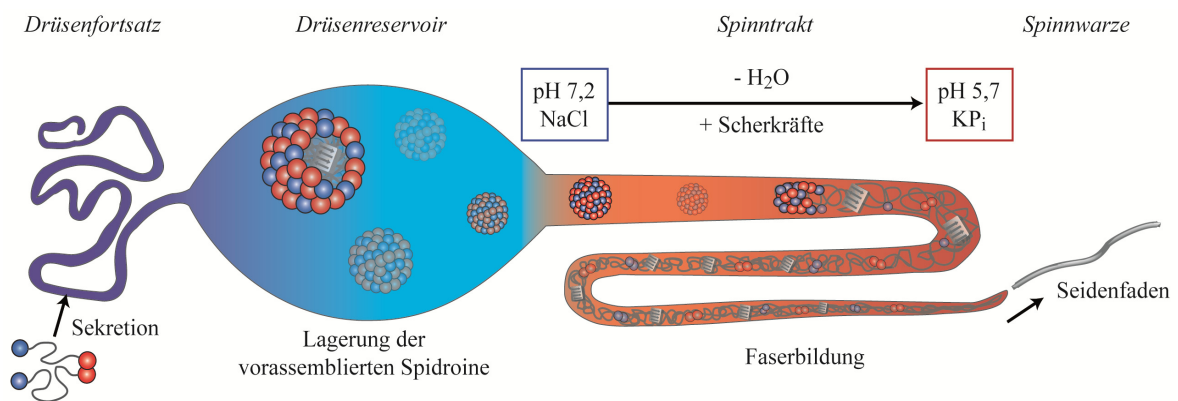
**Abbildung 2: Aufbau des Dragline-Fadens.** Der MA-Seidenfaden setzt sich aus fünf hierarchischen Schichten zusammen, die vereinfacht als Schale (Lipide, Glykoproteine, MI-Seide) und Kern (äußere und innere Region) kategorisiert werden. Die innerste Schicht enthält heterogene Proteinfasern mit kristallinen Bereichen (dicht-gepackte  $\beta$ -Faltblätter) und ist von einer amorphen Matrix (diverse Strukturelemente) umgeben. Modifiziert nach Heidebrecht *et al.*<sup>107</sup> mit Genehmigung des Verlages Elsevier.

Beide Kernregionen (Schicht 4 und 5) enthalten gestapelte Proteinfibrillen, jedoch unterscheiden sich die Fibrillen in den Mischungsverhältnissen von MaSp1 zu MaSp2 und den daraus resultierenden Strukturanteilen.<sup>102, 108</sup> Die äußere Kernregion besteht hauptsächlich aus dem Alanin-reichen MaSp1, das vorwiegend  $\beta$ -Faltblattstrukturen ausbildet und für die Starrheit der recht homogenen Schicht verantwortlich ist.<sup>109</sup> Die innere und eher heterogene Kernregion setzt sich aus einer Mixtur an MaSp1 und MaSp2 zusammen, die in semikristallinen Fibrillen entlang des Fadens organisiert sind.<sup>103</sup> Die von dem nahezu Prolin-freien MaSp1 ausgebildeten  $\beta$ -Faltblätter organisieren sich in dicht-gepackte, hochausgerichtete, kristalline Bereiche, die ca. 15 % des Gesamtvolumens des Dragline-Fadens ausmachen.<sup>110, 111</sup> Das laterale Wachstum der kristallinen Bereiche wird durch das Einbinden von MaSp2 begrenzt, dessen hoher Prolin-Gehalt (GPGXX-Motiv) die Bildung von dicht-gepackten  $\beta$ -Faltblattstrukturen unterdrückt und eine heterogene, amorphe Matrix aus nicht-ausgerichteten  $\beta$ -Faltblättern,  $\beta$ -Schleifen,  $\beta$ -Spiralen und  $3_{10}$ -Helices entstehen lässt.<sup>101, 110, 112-114</sup> Die außergewöhnlichen mechanischen Eigenschaften des Dragline-Fadens beruhen zu einem großen Teil auf dem semikristallinen Aufbau der zweiteiligen Kernregion, dessen kristalline Bereiche dem Faden ihre Zugfestigkeit verleihen und dessen amorphe Bereiche die Elastizität ermöglichen.<sup>115</sup> Allerdings kann die

Zusammensetzung der MA Seide deutlich zwischen verschiedenen Spinnenspezies und sogar zwischen den Seidenfäden einer einzigen Spinne variieren, da die Fadenbildung auch von zahlreichen Umwelteinflüssen wie beispielsweise der Ernährungsweise des Tieres abhängt.<sup>116, 117</sup>

### 1.6. NATÜRLICHER SPINNPROZESS

Die mechanischen Eigenschaften des *Dragline*-Fadens hängen neben der Zusammensetzung auch von dem zugrundeliegenden Assemblierungsmechanismus ab. Die Evolution der Spinndrüse hat einen hochkomplexen Spinnprozess entstehen lassen, der innerhalb weniger Millisekunden und unter physiologischen Bedingungen eine Proteinlösung zu einem mechanisch stabilen Seidenfaden transformiert.<sup>49, 56, 118</sup> Der komplexe Spinnapparat baut sich aus vier verschiedenen Hauptkomponenten auf (Abbildung 3).



**Abbildung 3: Schematische Darstellung des natürlichen Spinnprozesses.** Nach Sekretion im Spinnfortsatz werden die MaSp in Mizellen-ähnlichen Strukturen im Reservoir der Drüse gelagert. Im Spinntrakt induzieren chemische und mechanische Änderungen die Assemblierung der MaSp zu einem festen Faden, der final von der Spinne aus der Spinnwarze gezogen wird.

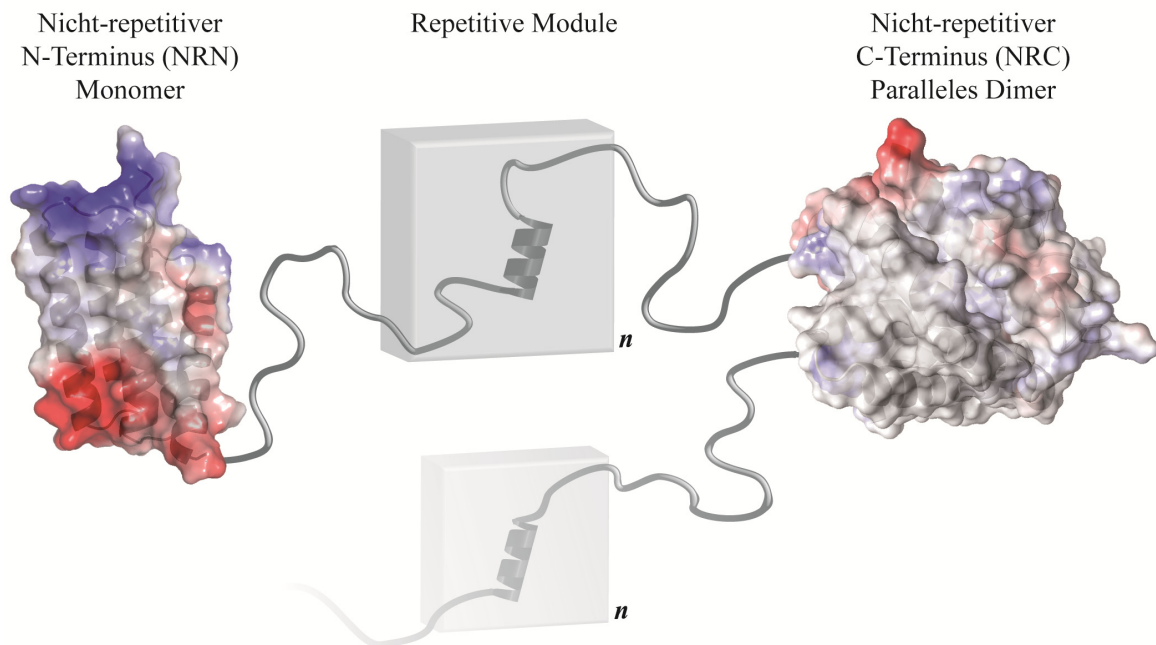
MaSp1 und MaSp2 werden zunächst simultan von speziellen Epithelzellen, die sich durch ein großes endoplasmatisches Retikulum und zahlreiche sekretorische Vesikel auszeichnen, in das Lumen des ungefähr 60 Millimeter langen, ausgestreckten Drüsenfortsatzes und des breiteren Drüsenreservoirs sekretiert.<sup>49, 101, 119-121</sup> Anschließend wird die auf bis zu 50 % (w/v) konzentrierte Proteinlösung bei neutralem pH-Wert und in der Anwesenheit von Natrium- und Chloridionen im Drüsenreservoir gelagert.<sup>122, 123</sup> Die amphiphilen Spidroine nehmen während der Lagerung in der Ampulla vermutlich Mizellen-ähnliche Strukturen ein, die gleichermaßen im Fall von Fibroin in der Drüse von *Bombyx mori* identifiziert wurden, um eine verfrühte und schädliche Aggregation zu verhindern.<sup>124-128</sup> In diesem Assemblierungsstadium bilden die präarrangierten Spidroine bereits lyotropische Flüssigkristalle aus.<sup>49, 67, 70, 129</sup> Anschließend werden die Mizellen-ähnlichen Assemblate aus

dem Reservoir durch die drei Windungen des konisch zulaufenden Spinntrakts geleitet. Die dabei auftretenden Elongations- und Scherkräfte forcieren die Koaleszenz der Strukturen und richten die  $\beta$ -Faltblattstrukturen der Flüssigkristalle weiter aus.<sup>129</sup> Der *Dragline*-Faden erhält seine typische Schale-Kern Struktur, indem das MaSp1-MaSp2 Spidroingemisch der Fibrillen-bildenden Kernregion mit einem homogenen Sekret aus Glykoproteinen beschichtet wird.<sup>49, 129</sup> Der Spinntrakt ist mit einer kutikularen Intima (Schicht aus Epithelzellen) beschichtet, die analog zu einer Hohlfaserdialysemembran die schnelle Resorption des Wassers (Dehydration) und den Austausch von Ionen ermöglicht.<sup>129</sup> Dabei werden die mäßig chaotropen Natrium- und Chloridionen durch vergleichsweise kosmotrope Kalium- und Phosphationen ausgetauscht, was durch das Aussalzen und Umstrukturieren der Spidroine den Übergang der Proteinlösung zu einem festen *Dragline*-Faden induziert.<sup>123, 130</sup> Zudem senken luminale Carboanhydrasen und epitheliale Ionenkanäle den pH-Wert der Spinnlösung von pH 7,2 auf mindestens pH 5,7 ab.<sup>123, 130-133</sup> Die Ansäuerung und Dehydration verstärken die vorwiegend in  $\beta$ -Faltblattstrukturen vorkommenden Wasserstoffbrückenbindungen und lösen Änderungen in der Tertiär- und Quartärstruktur der terminalen Spidroindomänen aus.<sup>134</sup> Der natürliche Spinnprozess wird von der Spinne abgeschlossen, indem der Seidenfaden unter Strecken aus der Spinnwarze gezogen wird. Das ventilähnliche Organ wird von Muskeln gesteuert und für die Wiederaufnahme des Spinnprozesses benötigt.<sup>129</sup>

## 1.7.SPINNENSEIDENPROTEINE

### 1.7.1 NATÜRLICHE SPINNENSEIDENPROTEINE

Obwohl es eine große Vielfalt an verschiedenen Spinnenspezies und Seidentypen gibt, weisen die meisten zugrundeliegenden Spidroine eine konservierte Komposition aus drei verschiedenen Domänen auf (Abbildung 4).<sup>135</sup> Die Spidroine des *Dragline* Fadens enthalten eine bis zu 350 kDa große Kerndomäne aus bis zu hundertfach wiederholten, spezifischen Sequenzmotiven.<sup>41, 75, 82, 101, 136, 137</sup> Die Art, Anzahl und Reihenfolge der repetitiven Module variiert deutlich zwischen den verschiedenen Seidenarten, um die mechanischen Eigenschaften des jeweiligen Seidenfadens an die spezifische ökologische Funktion anzupassen.<sup>54, 98, 99</sup> Die Kerndomäne wird an dem N- und C-Terminus durch hochkonservierte, globulär gefaltete und nicht-repetitive amino- (NRN) bzw. carboxy-terminale Domänen (NRC) komplettiert, die sowohl die Lagerung der Spidroine in löslicher Form als auch die Assemblierung zu einem stabilen Faden unterstützen.<sup>138, 139</sup>



**Abbildung 4: Schematischer Aufbau der löslichen Form natürlicher Spidroine.** Die Spidroine bestehen aus einer großen ( $\leq 350$  kDa), ungefalteten Kerndomäne mit  $n$ -Wiederholungen der repetitiven Module, die von kleinen ( $< 15$  kDa), globulär gefalteten terminalen Domänen flankiert wird. Während die nicht-repetitive N-terminale Domäne (NRN) bei neutralem pH-Wert den monomeren Zustand einnimmt, liegt die nicht-repetitive C-terminale Domäne (NRC) in den meisten MaSp als kovalent gebundenes Dimer vor. Die NRN (Struktur aus Hagn *et al.*<sup>140</sup>) und NRC Domäne (PDB ID: 2KHM) wurde mit PyMOL<sup>141</sup> dargestellt.

#### 1.7.2 REKOMBINANT HERGESTELLTE SPINNENSEIDENPROTEINE

Um die Rolle der Spidroindomänen bei der Selbstassemblierung zu untersuchen, werden diese einzeln oder kombiniert in verschiedenen Wirtsorganismen wie beispielsweise Bakterien (*Escherichia coli*),<sup>142</sup> Hefen (*Pichia pastoris*),<sup>143</sup> Insekten (Seidenraupenlarve),<sup>144-146</sup> Säugetierzellen (Hamster),<sup>147</sup> Pflanzen (Tabak, Kartoffel, Sojabohne, *Arabidopsis*)<sup>148, 149</sup> und transgenen Tiere (Maus, Ziege)<sup>150-152</sup> rekombinant hergestellt.


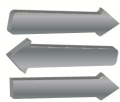


Die Produktion natürlicher MaSp-Sequenzen aus cDNA (englisch *complementary deoxyribonucleic acid*) in transgenen Organismen und Zellen erwies sich auf Grund der hohen Kosten, der komplizierten Aufzucht und der niedrigen Ausbeute (Maus:  $12 \text{ mg L}^{-1}$ , Insektenzelle:  $50 \text{ mg L}^{-1}$ , Säugetierzelle:  $50 \text{ mg L}^{-1}$ , Ziege: keine Angabe)<sup>142, 147, 151, 153-155</sup> als unrentable.<sup>156</sup> Dabei beschränkte insbesondere der signifikant unterschiedliche Codongebrauch zwischen den Spinnen und dem Wirtsorganismus sowie die repetitive Natur der Seidengene die Verwendung natürlicher Gensequenzen.<sup>157</sup>

Diese Schwierigkeiten wurden überwunden, indem aus den Motivwiederholungen der natürlichen Gensequenz ein an den Codongebrauch des Wirtsorganismus angepasstes Konsensusmotiv abgeleitet wurde, dessen Multimerisierung die Herstellung rekombinanter

Seidenproteine mit variierendem Molekulargewicht von bis zu 250 kDa ermöglichte.<sup>158, 159</sup> Auf diese Weise wurden auf den Konsensussequenzen von *Araneus diadematus* Fibroin 3 (ADF3) und 4 (ADF4) basierende Module entworfen (englisch *engineered*, eADF3 und eADF4), die beliebig miteinander kombiniert werden können.<sup>160</sup> Die NRN- und NRC-codierenden Sequenzen verschiedener Spidroine können einzeln exprimiert, mit Hilfe von ortsspezifischer Mutagenese modifiziert und mit den Sequenzen von eADF3 und eADF4 fusioniert werden.<sup>107, 138, 139</sup> Die konstruierte Gensequenz wird anschließend in einen Expressionsvektor inseriert und in die Wirtszelle transformiert. Mit Hilfe etablierter Fermentationsprozesse kann der Wirtsorganismus kultiviert und das rekombinante Protein produziert werden.<sup>156</sup> Anschließend wird das rekombinante Spidroin durch eine Reinigungsstrategie aus hitze- und salzinduzierten Fällungen sowie chromatographischen Trennverfahren von den wirtseigenen Proteinen abgetrennt<sup>160</sup> und steht für analytische Untersuchungen und Anwendungen bereit. Dabei produzieren Hefen die rekombinanten Spidroine mit einer Ausbeute von bis zu 2,9 g L<sup>-1</sup>.<sup>161</sup> In den meisten Fällen wird jedoch das etablierte *E. coli*-Expressionssystem zur Herstellung der rekombinanten Spidroine verwendet, da dieser Ansatz eine einfache Manipulation des Wirtsorganismus, eine kurze Generationszeit, niedrige Produktionskosten, die Möglichkeit zum *upscaling* (englisch, Hochskalierung) der Produktion, eine hohe Ausbeute (bis zu 1,2 g L<sup>-1</sup>)<sup>159</sup> und eine gleichbleibende Qualität miteinander vereint.<sup>156</sup>

### 1.7.3 KERNDOMÄNE

Ein Vergleich der Aminosäuresequenzen natürlicher und rekombinanter Spidroine ermöglicht, aus den repetitiven Aminosäuresequenzen spezifische Proteinstrukturmotive abzuleiten, mit denen die in der Spinnenseide vereinte Kombination aus Zugfestigkeit und Elastizität erklärt werden kann. Die sich wiederholenden Aminosäuresequenzen werden im Folgenden anhand von FLAG (Flagelliform-Spidroin), MiSp, MaSp1 und MaSp2 aus *Nephila clavipes*, Fibroin 3 und 4 (ADF3 und ADF4, beides MaSp2) aus *Araneus diadematus* sowie den auf der Konsensussequenz von ADF3 und ADF4 basierenden, rekombinanten Modulen A und Q von eADF3 (*engineered*, englisch entworfen) und dem Modul C von eADF4 diskutiert. Die repetitiven Peptidmotive können in die vier Hauptkategorien (1) GPGXX, (2) Alanin-reich, (3) GGX und (4) *Spacer* (englisch, Abstandshalter) eingeteilt werden (Abbildung 5), wobei die Abkürzung G für Glycin, P für Prolin und X für eine beliebige Aminosäure steht.<sup>39, 41, 62, 75, 82, 101, 162, 163</sup>

		 β-Schleife/Spirale GPGXX <i>elastisch</i>	 β-Faltblatt Alanin-reich <i>kristallin</i>	 3 <sub>10</sub> -Helix GGX <i>amorph</i>	 Spacer
Spidroin Modul					
<i>N. clavipes</i>	FLAG	✓		✓	✓
	MiSp		✓	✓	✓
	MaSp1		✓	✓	
	MaSp2	✓	✓		
<i>A. diadematus</i>	ADF3	✓	✓	✓	
	ADF4	✓	✓		
Rekombinant	eADF3		✓	✓	
	eADF3	✓			
	eADF4	✓	✓		

**Abbildung 5: Struktur motive der Kerndomänen natürlicher und rekombinanter Spidroine.** Das G der Aminosäuremotive steht für Glycin, P für Prolin und X für eine variierende Aminosäure, die in den meisten Fällen durch ein Alanin, Glutamin, Glycin, Leucin, Serin oder Tyrosin eingenommen wird (angelehnt an Scheibel *et al.*<sup>164</sup> und Heidebrecht *et al.*<sup>107</sup>). FLAG, Flagelliform-Spidroin; MiSp, minor ampullate spidroin; MaSp, major ampullate spidroin; ADF, *Araneus diadematus* fibroin; eADF, rekombinantes ADF bestehend aus den Modulen A, Q und C.

Das Sequenzmotiv GPGXX (1) kann weiter in die Kategorien GPGGX und GPGQQ aufgeteilt werden. Die Ausbildung von klassischen Strukturelementen wird durch das im Motiv enthaltene Prolin unterbrochen, da die Seitenkette dieser Aminosäure eine zyklische Struktur einnimmt und das Stickstoffatom somit nicht als Protonendonator agieren kann, um die Sekundärstruktur über Wasserstoffbrückenbindungen zu stabilisieren.<sup>165</sup> Als Folge dessen nehmen die GPGXX-Sequenzen bevorzugt eine β-Spiral-Konformation ein, die aus der multiplen Aneinanderreihung von β-Schleifen des Typs II entsteht.<sup>75, 166</sup> Die zahlreichen Proline lassen die Polypeptidkette mehrfach abknicken und eine flexible Spiralstruktur entstehen, die in der Kernregion des Seidenfadens als amorphe Matrix dient und dem Faden



seine Elastizität verleiht.<sup>163</sup> Infolgedessen ist das GPGXX-Motiv besonders in den stark dehnbaren Seidentypen Flagelliform und MaSp (MaSp2, ADF3, ADF4) anzutreffen<sup>163</sup> und wurde in den Konsensussequenzen der rekombinanten Spidroine eADF3 (Modul Q) und eADF4 inkludiert.<sup>160</sup> Interessanterweise korreliert die Anzahl der Motivwiederholungen in MaSp2 (bis zu neun aneinandergereihte  $\beta$ -Schleifen) und in dem Flagelliform-Spidroin (mindestens 43 zusammenhängende  $\beta$ -Schleifen) direkt mit der Dehnbarkeit des resultierenden *Dragline* (35 %) und Flagelliform-Seidenfadens (200 %).<sup>51, 75, 101</sup>

Die in MiSp, MaSp und ADF auftretende Kategorie der Alanin-reichen Motive (2) umfasst Poly-Alanin und Poly-GA Bereiche, die bevorzugt parallel zur Faserachse ausgerichtete  $\beta$ -Faltblattstrukturen einnehmen.<sup>40, 166-171</sup> Dabei interagieren die benachbarten Proteinstränge über die hydrophoben Alanin-Seitenketten und ermöglichen eine dichte Stapelung der  $\beta$ -Faltblätter.<sup>163</sup> Die resultierenden kristallinen Bereiche verleihen den MaSp- (1,2 GPa, *Araneus diadematus*) und MiSp-Seidenfäden (1,4 GPa, *Araneus gemmoides*) ihre hohe Zugfestigkeit, wohingegen der Flagelliform-Seidenfaden auf Grund der fehlenden  $\beta$ -Faltblattstrukturen deutlich weniger zugfest ist (0,5 GPa, *Araneus diadematus*).<sup>49, 54, 55, 70, 99, 172</sup> Die rekombinanten Spidroine eADF3, die einen (Ala)<sub>6</sub>-Bereich in der Konsensussequenz des Moduls A tragen, wurden erfolgreich zu artifiziellen Seidenfasern mit einer zu natürlichen *Dragline*-Fäden vergleichbaren Zähigkeit versponnen.<sup>160, 173</sup>

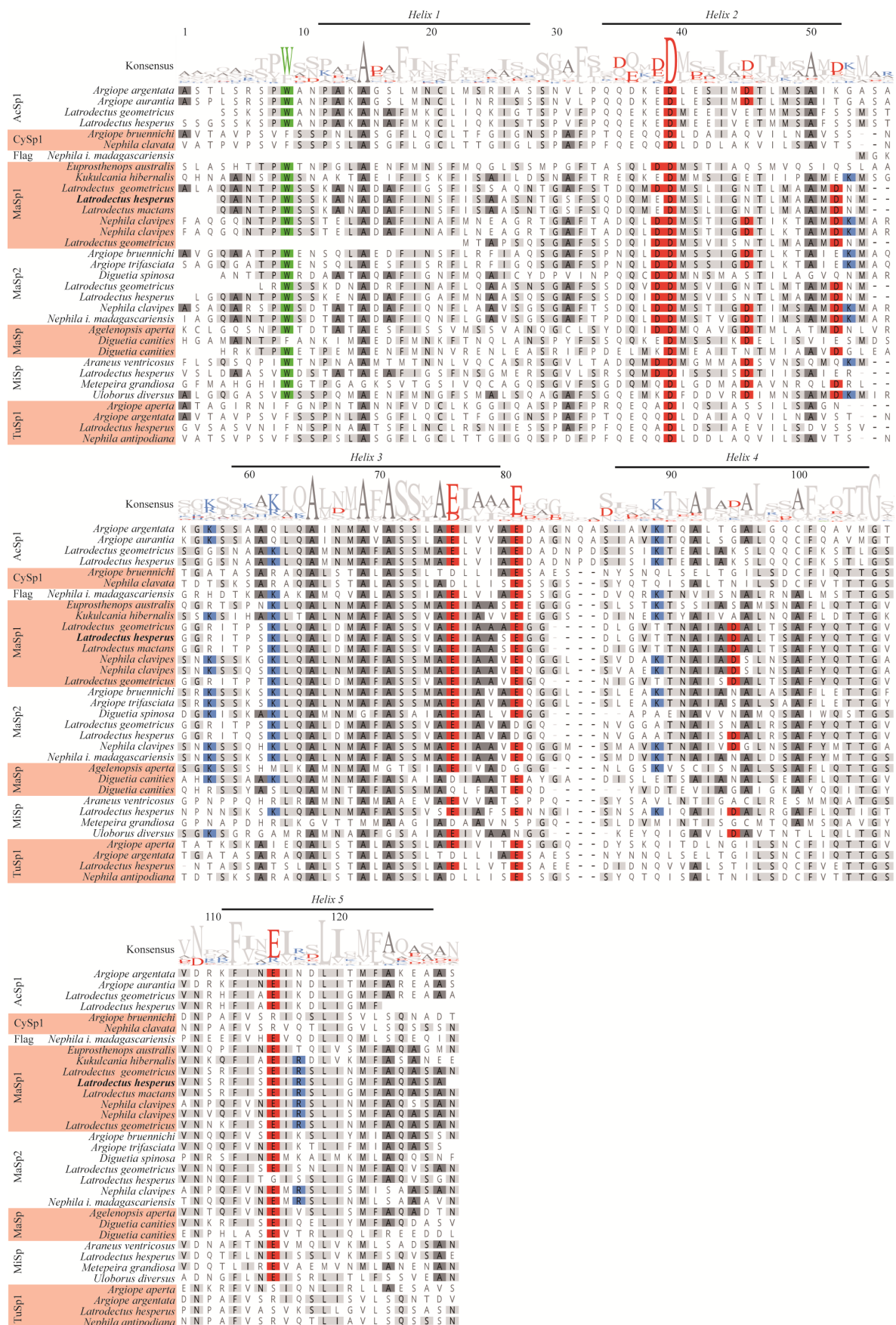
Das GGX-Motiv (3) aus FLAG, MiSp, MaSp1, ADF3 und eADF3 (Modul A) favorisiert eine  $3_{10}$ -Helix-Konformation, bei der sich die benachbarten Strukturmodule über interhelikale Wasserstoffbrückenbindungen quervernetzen und einheitlich ausrichten.<sup>166, 170, 174, 175</sup> Zusätzlich können die GGX-Sequenzen in  $\beta$ -Schleifen bzw.  $\beta$ -Helices vorliegen.<sup>114, 175, 176</sup> Die reversible Faltung dieser Strukturen lässt eine flexible und amorphe Matrix entstehen, in die die Kristalle der Alanin-reichen Bereiche eingebettet werden.<sup>163</sup>

Die repetitiven Einheiten der MiSp und Flagelliform-Seidenproteine werden durch nicht-repetitive und Serin-haltige *Spacer* unterbrochen.<sup>82</sup> Die exakte Funktion der *Spacer* konnte bisher nicht determiniert werden. Es wird jedoch angenommen, dass die *Spacer* die Spidroine während der Lagerung eine alternative Faltung einnehmen lassen, um deren Löslichkeit zu erhöhen. Während der Fadenbildung scheinen die *Spacer* die Ausrichtung der geordneten Strukturelemente, insbesondere die der kristallinen Bereiche, zu unterstützen und eine Kontaktfläche zu andersartigen Komponenten (z.B. der Aggregatseide) zu bieten.<sup>163</sup>

#### 1.7.4 AMINO-TERMINALE DOMÄNE

Die nicht-repetitive N-terminale Domäne (NRN) ist der am höchsten konservierte Sequenzbereich und weist zwischen homologen Spidroinen verschiedener Spinnenspezies und Seidentypen eine sehr ähnliche Aminosäuresequenz und Ladungsverteilung auf (Abbildung 6). Dabei sind die sauren Aminosäuren an den Positionen D39, E76, E81 und E114 besonders stark konserviert, was deren Bedeutung für die Faltung und/oder Funktion der Domäne hervorhebt.<sup>139, 177-181</sup> Zur besseren Vergleichbarkeit der NRN verschiedener Spidroine werden in dieser Arbeit die Aminosäurepositionen einheitlich nach der Sequenz von NRN1<sub>L.h.</sub> aus *Latrodectus hesperus* MaSp1 nummeriert.

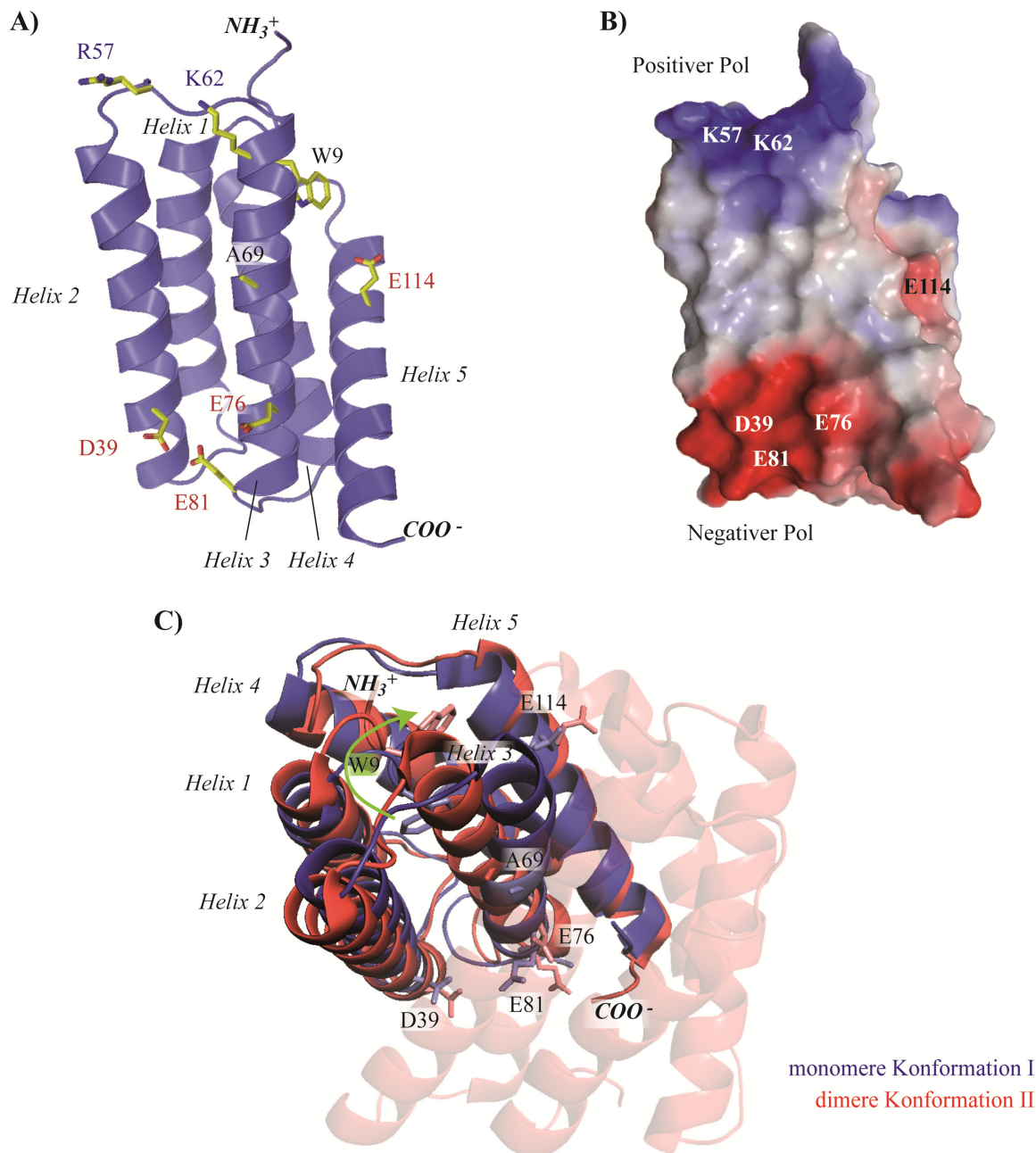
Der N-terminale Genbereich von NRN enthält multiple Startcodons (Methionin), die vermutlich während der Translation mehrfache Initiierungsmöglichkeiten für die Assemblierung des Ribosoms bieten. Die zusätzliche Synthese von kürzeren Spidroinisoformen ermöglicht die für den Spinnprozess benötigte Proteinkonzentration zu erreichen.<sup>177, 179</sup> Die NRN Domäne ist zudem N-terminal mit einem Signalpeptid modifiziert (in Abbildung 6 nicht aufgeführt), welches das Spidroin in das endoplasmatische Retikulum lenkt, die Sekretion via Golgi Apparat in das Lumen der Spinndrüse ermöglicht und anschließend über spezifische Schnittstellen abgespalten wird<sup>177, 182</sup>



**Abbildung 6: Sequenzalignment der N-terminalen Domänen verschiedener Spidroine.** Die Aminosäuresequenzen der N-terminalen Domänen 33 verschiedener Spidroine wurden mittels Geneious<sup>183</sup> abgeglichen. Aminosäuren wurden ab einer Identität von 25 % in der Konsensussequenz berücksichtigt (hellgrau hinterlegt, Tryptophan grün, Alanin dunkelgrau,

saure Aminosäuren rot und basische blau). Die Größe der Aminosäurebuchstaben in der Konsensussequenz entspricht der prozentualen Identität der Aminosäure. Mit Hilfe von Jpred 4<sup>184</sup> wurden die helikalen Bereiche der Konsensussequenz berechnet. AcSp1, Aciniform spidroin 1; CySp1, Cylindriförmiges spidroin 1; Flag, Flagelliform spidroin; MaSp1, Major Ampullate Spidroin 1; MaSp2, Major Ampullate Spidroin 2; MiSp, Minor Ampullate Spidroin; TuSp1, Tubuliform Spidroin 1.

Im Vergleich zu den repetitiven Modulen der Kerndomäne weist die ungefähr 130 Aminosäuren lange und 14 kDa große NRN Domäne eine vollständig unterschiedliche Aminosäuresequenz und damit korrelierend eine alternative Faltung auf. Mit Hilfe von Kristallstrukturanalysen bzw. Kernspinresonanz-Spektroskopie (englisch *nuclear magnetic resonance*, NMR) wurden für die homologen NRN Domänen aus *Euprosthenops australis* (NRN<sub>E.a.</sub>) und *Latrodectus hesperus* MaSp1 (NRN<sub>L.h.</sub>) eine homologe Faltung in ein Fünf-Helix-Bündel gezeigt (Abbildung 7A).<sup>139, 140, 178, 185</sup> Dabei sind die hochkonservierten Ladungen auf der Oberfläche exponiert und in einem negativ geladenen *Clusters* (D39, E76 und E81; E114 liegt peripher) und einem positiv geladenen *Cluster* (R57 und K62) organisiert. Die Ladungsverteilung der Proteinoberfläche lässt einen makroskopischen Dipol entstehen (Abbildung 7B).<sup>139, 140, 185</sup> Das Fünf-Helix-Bündel liegt bei neutralem pH Wert und in der Anwesenheit von Natrium- und Chlorid-Ionen (NaCl) als stabiles Monomer vor und unterdrückt dabei die Aggregation der Kerndomäne, wie anhand von rekombinanten Mini-Spidroinen (27 kDa) gezeigt wurde.<sup>139</sup> Sobald der pH-Wert und die NaCl-Konzentration entlang des Spinntrakts abfallen,<sup>123, 130-133</sup> konvertiert die Tertiärstruktur innerhalb von lediglich 60 µs von der lockeren Konformation I in die dichter gepackte Konformation II (Abbildung 7C).<sup>131, 139, 185-188</sup> Während der Strukturänderung verdreht sich eine Einheit aus Helix 2, die die konservierte Aminosäure D39 trägt, und Helix 3, auf der E76 und E81 lokalisiert sind, um ca. 5 Å und 16° um die starre Gruppe aus den Helices 1, 4 und 5.<sup>185</sup> Gleichzeitig wird das einzige Tryptophan (W9) aus dem hydrophoben Inneren des Fünf-Helix-Bündels zur hydrophileren Oberfläche verschoben, was anhand eines bathochrom (rot-) verschobenen Fluoreszenzmaximums und einer reduzierten Emissionsintensität spektroskopisch detektiert werden kann.<sup>139, 140, 185, 188</sup> Simultan zu der strukturellen Änderung bildet NRN im sauren pH Bereich über die Oberfläche der Helices 2, 3 und 5 antiparallele Dimere aus, was die Selbstassemblierung der Kerndomäne beschleunigt.<sup>139, 140</sup> Die überdurchschnittlich hohe Assoziationskonstante von  $k_a = 3 \times 10^8 \text{ M}^{-1} \text{ s}^{-1}$  (pH 6,0) deutet auf einen effektiven Dimerisierungsmechanismus hin,<sup>189</sup> der es ermöglicht die Spidroinassemblierung innerhalb von Millisekunden ablaufen zu lassen, um eine Spinngeschwindigkeit von 60-400 mm s<sup>-1</sup> zu erreichen.<sup>56, 190</sup> Es wird angenommen, dass die im sauren pH Bereich ausgelöste Konformationsänderung und Dimerisierung über die Protonierung der hochkonservierten, sauren Aminosäureseitenketten gesteuert wird.



**Abbildung 7: Struktur der N-terminalen Domäne (NRN).** **A)** Die konservierte Faltung von NRN in eine Fünf-Helix-Bündel Struktur wird am Beispiel von  $\text{NRN}_{L.h.}$  aus *L. hesperus* MaSp1 dargestellt (kein PDB-Eintrag)<sup>140</sup>. Die Seitenketten der besonders konservierten Aminosäuren W9, D39, R57, A69, K62, E76, E81 und E114 wurden skizziert. Farbmarkierung der Atome: C, gelb; N, blau; O, rot. **B)** Die Ladungsverteilung des wie in A) orientierten  $\text{NRN}_{L.h.}$  gleicht einem makroskopischen Dipol, da die sauren bzw. basischen Aminosäuren lokal in einem negativ bzw. positiv geladenen *Cluster* kumulieren. Gezeigt wird das elektrostatische Oberflächenpotential von  $+5 \text{ kT e}^{-1}$  (blau) bis  $-5 \text{ kT e}^{-1}$  (rot). **C)** Bei neutralem pH Wert und in der Anwesenheit von 300 mM NaCl nimmt  $\text{NRN}_{E.a.}$  aus *Euprosthenops australis* MaSp1 (PDB ID: 2LPJ)<sup>185</sup> Konformation I (blau) ein, bei der das W9 im Inneren des Fünf-Helix-Bündels verborgen ist. In saurer (pH 5,5) und NaCl-freier Umgebung liegt  $\text{NRN}_{E.a.}$  (PDB ID: 2LTH)<sup>191</sup> in der dichter gepackten Konformation II (rot) vor. Dabei haben sich die im Dimerisierungsinterface lokalisierten Helices 2/3 um die starre Einheit aus Helices 1/4/5 gedreht, wodurch das W9 aus dem Proteininneren zur Oberfläche verschoben wird. Die zweite Untereinheit des bei pH 5,5 vorliegenden Dimers wurde in einem transparenten Rotton dargestellt. Die Abbildungen wurden mit PyMOL erstellt.<sup>141</sup>

Die in dem Protonengradient zwischen pH 7,2 und pH 5,7 ablaufenden molekularen Mechanismen der Konformationsänderung und Dimerisierung rückten in den Fokus zahlreicher Untersuchungen. Da die Seitenkette eines freien Histidins einen  $pK_S^{\text{frei}}$  von 6,0 aufweist, wurde zunächst angenommen, dass die Protonierung vereinzelt vorkommender Histidine die Änderungen in der Tertiär- und Quartärstruktur auslöst.<sup>192, 193</sup> Im monomeren NRN<sub>E.a.</sub> scheint die Interaktion des H6 mit benachbarten Ladungen und Dipolen<sup>194-196</sup> den  $pK_S^{\text{H6}}$  auf 6,3 - 6,8 anzuheben (Tabelle 2). Die Deletion des nicht-konservierten H6 beeinflusste jedoch nicht die Eigenschaften von NRN<sub>E.a.</sub>,<sup>139</sup> sodass anschließend trotz deutlich niedrigerer  $pK_S$ -Werte (Tabelle 2) die Rolle der stark konservierten Asparagin- und Glutaminsäuren (Abbildung 6) untersucht wurde. Die  $pK_S$ -Werte werden offensichtlich durch die lokale Kumulation der Carboxylgruppen der sauren Aminosäuren im negativen Pol leicht angehoben,<sup>194-196</sup> sodass die Protonierungen bzw. Neutralisierungen der Carboxylatgruppen zwischen pH 7,2 und 5,5 ermöglicht und die energetisch ungünstigen coulombsche Repulsionen im negativen *Cluster* abgeschwächt werden.<sup>189</sup> An NRN<sub>E.a.</sub> durchgeführte Simulationen offenbarten leicht unterschiedliche  $pK_S$ -Werte,<sup>197, 198</sup> jedoch wird die Aussagekraft der Berechnungen durch die Verwendung einer mit Artefakten behafteten Kristallstruktur (PDB ID: 3LR2) eingeschränkt, da die  $pK_S$ -Werte insbesondere durch die Fehlausrichtung der Seitenkette E81 in die unmittelbare Nähe der Seitenkette D39 verfälscht werden.<sup>191</sup>

**Tabelle 2:  $pK_S$ -Werte von individuellen und in der Struktur des monomeren NRN<sub>E.a.</sub> vorkommenden Aminosäureseitenketten.** Neben den  $pK_S$ -Werten der Seitenketten von freien Aminosäuren sind die mit verschiedenen Methoden berechneten  $pK_S$ -Werte der hochkonservierten Aminosäureseitenketten von NRN<sub>E.a.</sub> angegeben.

Freie Aminosäure <sup>192, 193</sup>		Anandakrishnan <i>et al.</i> <sup>199</sup>		Wallace & Shen <sup>197</sup>	da Silva <i>et al.</i> <sup>198</sup>
		PDB	2LPJ (NMR)	3LR2 (X-ray)	3LR2 (X-ray)
H	6,0	H6	6,5	6,8	6,3
D	4,0	D39	4,0	4,1	4,3
		E76	4,3	4,4	4,6
E	4,4	E81	4,5	4,4	4,7
		E114	3,8	4,2	4,4

Die spezifische Funktion einzelner konservierter Aspartate und Glutamate wurde in einigen Veröffentlichungen untersucht, indem die selektive Protonierung individueller Seitenketten mit Hilfe von gentechnischer Manipulation imitiert wurde. Dabei wurden die Asparagin- und Glutaminsäuren von NRN<sub>E.a.</sub>,<sup>139, 185, 187, 189</sup> NRN<sub>A.v.</sub> aus *Araneus ventricosus* MiSp<sup>180</sup> und

NRN<sub>N.c.</sub> aus *Nephila clavipes* MaSp1<sup>200</sup> mittels ortsspezifischer Mutagenese durch die nicht-titrierbaren und neutralen Analoga Asparagin und Glutamin ersetzt. Die Varianten wurden insbesondere mit Hilfe der pH-abhängigen Fluoreszenzänderung des W9 charakterisiert.

In der Wildtyp-Domäne laufen die Dimerbildung und die Strukturänderung nahezu parallel zwischen pH 6,4 und pH 6,2 ab.<sup>139</sup> Da keiner der erstellten Datensätze über mutierte NRN die Erkenntnisse über die Strukturänderung mit Daten über das Monomer-Dimer Gleichgewicht vereint, war der Zusammenhang zwischen den beiden Prozessen (Voraussetzung, Folge, Unabhängigkeit) nicht bekannt. Die nicht erwiesene Korrelation beider Prozesse wurde jedoch in den darauf folgenden Veröffentlichungen vorausgesetzt, indem das Monomer-Dimer Gleichgewicht der Varianten von NRN<sub>E.a.</sub>, NRN<sub>A.v.</sub> und NRN<sub>N.c.</sub> nicht direkt analysiert sondern indirekt über die Veränderungen der Trp-Fluoreszenz beurteilt wurde.<sup>180, 191, 200</sup> Als Folge dessen sollten die im Folgenden zitierten Schlussfolgerungen über die Rolle individueller Aminosäuren während der NRN Dimerisierung kritisch betrachtet werden (Tabelle 3).

Der Seitenkette der D39 wird eine Schlüsselrolle für die Kontrolle der Strukturänderung zugesprochen, da in der Variante D39N der bathochrome Effekt der Trp-Fluoreszenz bei deutlich niedrigerem pH-Wert (< pH 6) auftritt.<sup>139, 191</sup> Dabei scheint D39 die pK<sub>S</sub>-Werte der Carboxylgruppen der benachbarten E76 und E81 im sauren *Cluster* anzuheben (Tabelle 2) bzw. deren Protonierung bei höheren pH Werten zu ermöglichen. Gleichzeitig verhindert der auffällig niedrigere pK<sub>S</sub>-Wert von D39 (Tabelle 2) die Aufnahme eines Protons während der Dimerbildung.<sup>198</sup> Somit kann die permanente Ladung von D39 das Dimer über eine intermolekulare Salzbrücke mit R57 bzw. K62 (Abstand D39–K62 < 4 Å)<sup>201</sup> aus dem relativ konservierten (Abbildung 6) positiven *Cluster* der zweiten Untereinheit stabilisieren.<sup>139, 189, 191, 197, 200-202</sup> Da die Energiebarriere der Salzbrücke D39–K62 in der Anwesenheit von NaCl um 29 % steigt, scheint die Asparaginsäure zudem an der Salz-Sensitivität der NRN-Dimerisierung beteiligt zu sein.<sup>201</sup> Umstrittene Hypothesen reichen von der Annahme, dass die Protonierung von D39 bei niedrigerem pH-Wert die Bindungsenergie des Dimers nicht beeinflusst,<sup>198</sup> bis hin zu der Theorie, dass das resultierende Aufbrechen der Salzbrücke die Dimerbildung vollständig verhindert.<sup>187, 189</sup>

**Tabelle 3: Zusammenfassung der in der Literatur beschriebenen Rollen der konservierten Asparagin- und Glutaminsäuren von NRN.** In den Referenzen wurde der Einfluss der Protonierung auf die Struktur experimentell bestimmt. Im Gegensatz dazu wurde das Monomer-Dimer Gleichgewicht der Mutanten lediglich anhand von Strukturdaten vorhergesagt, weshalb die aufgestellten Hypothesen über die Dimerisierung experimentell verifiziert werden müssen. AS, Aminosäure.

AS	Molekulare Interaktion	Konformation	Hypothetisches Monomer-Dimer Gleichgewicht	Offene Fragen
<b>D39</b>	<ul style="list-style-type: none"> <li>• Senkt <math>pK_s</math> der benachbarten AS (E79, E81)</li> <li>• Salzbrücke mit R57 und K62</li> </ul>	<ul style="list-style-type: none"> <li>• Carboxylatgruppe unterstützt Übergang von Konformation I zu II</li> </ul>	<ul style="list-style-type: none"> <li>• Negative Ladung der Carboxylatgruppe stabilisiert Dimer über Salzbrücke</li> </ul>	<ul style="list-style-type: none"> <li>• Zusammenspiel mit anderen AS</li> <li>• Einfluss auf Dimerisierung</li> </ul>
<b>E76</b>	<ul style="list-style-type: none"> <li>• Divergente Ausrichtung der Seitenkette zwischen verschiedenen NRN</li> </ul>	<ul style="list-style-type: none"> <li>• Selektive Protonierung hat keinen Einfluss</li> <li>• Gemeinsame Protonierung mit E114 stabilisiert Konformation II</li> </ul>	<ul style="list-style-type: none"> <li>• Umstritten</li> </ul>	<ul style="list-style-type: none"> <li>• Einfluss auf Dimerisierung</li> </ul>
<b>E81</b>	<ul style="list-style-type: none"> <li>• <i>Handshake</i> mit D39 widerlegt</li> <li>• Salzbrücke mit R57 und K62</li> </ul>	<ul style="list-style-type: none"> <li>• Protonierung hat Einfluss auf Konformation II</li> </ul>	<ul style="list-style-type: none"> <li>• Alleinige Protonierung destabilisiert Dimer</li> <li>• Gemeinsame Protonierung mit E76 und E114 stabilisiert Dimer</li> </ul>	<ul style="list-style-type: none"> <li>• Zusammenspiel mit anderen AS</li> <li>• Einfluss auf Dimerisierung</li> </ul>
<b>E114</b>	<ul style="list-style-type: none"> <li>• Abseits des negativen <i>Clusters</i> positioniert</li> </ul>	<ul style="list-style-type: none"> <li>• Selektive Protonierung hat keinen Einfluss</li> <li>• Gemeinsame Protonierung mit E76 stabilisiert Konformation II</li> </ul>	<ul style="list-style-type: none"> <li>• Gemeinsame Protonierung mit E76 stabilisiert Dimer</li> </ul>	<ul style="list-style-type: none"> <li>• Einfluss auf Dimerisierung</li> </ul>
<b>A69</b>	<ul style="list-style-type: none"> <li>• Interaktion mit A69 der zweiten Untereinheit</li> </ul>	<ul style="list-style-type: none"> <li>• A69R stabilisiert Konformation I</li> </ul>	<ul style="list-style-type: none"> <li>• A69R liegt als stabiles Monomer vor</li> </ul>	<ul style="list-style-type: none"> <li>• Zusammenhang Konformation und</li> </ul>



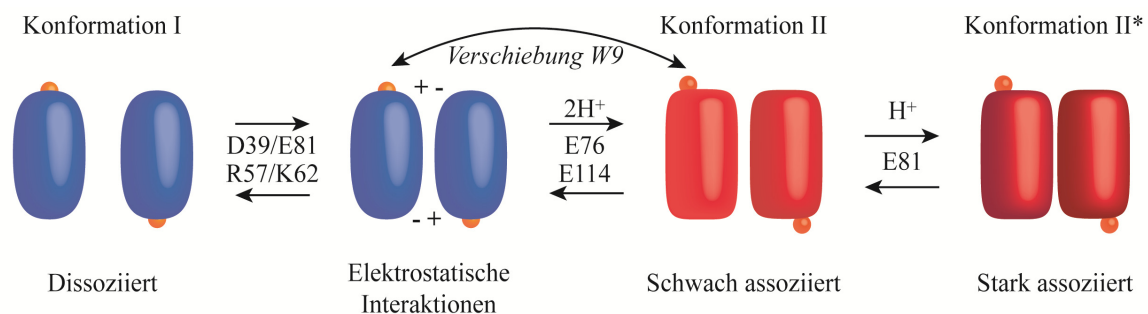
Die Ausrichtung der E76 unterscheidet sich leicht zwischen den gelösten Strukturen von  $\text{NRN}_{E.a.}$ ,  $\text{NRN}_{A.v.}$  und  $\text{NRN}_{N.c.}$ .<sup>139, 180, 200</sup> Da die Protonierung der Carboxylatgruppe die Struktur nicht verändert,<sup>139</sup> ist der Einfluss der Seitenkette auf das Monomer-Dimer Gleichgewicht umstritten. Für die Variante E76Q aus  $\text{NRN}_{E.a.}$  wurde neben einer wildtypischen, pH-abhängigen Dimerisierung<sup>189, 202</sup> sowohl die Stabilisierung der Tertiärstruktur im monomeren,<sup>198</sup> als auch im dimeren Zustand berichtet.<sup>187, 191, 197</sup>

Die im sauren *Cluster* lokalisierte E81 schien in der Kristallstruktur von  $\text{NRN}_{E.a.}$  eine *Handshake*-Interaktion (englisch, Seitenketten sind wie bei einem Handschlag zueinander ausgerichtet) mit D39 auszuüben,<sup>139</sup> allerdings wurde dies später in einer mittels NMR Spektroskopie gelösten Struktur widerlegt, da die Distanz zwischen den beiden Seitenketten  $> 8 \text{ \AA}$  betrug.<sup>191</sup> Verglichen mit  $\text{NRN}_{E.a.}$  ist E81 in  $\text{NRN}_{N.c.}$  unterschiedlich ausgerichtet und in  $\text{NRN}_{A.v.}$  ganz absent (Abbildung 6), sodass sie von einer Glutaminsäure an anderer Position vertreten wird.<sup>180, 200</sup> Während die selektive Neutralisierung der Glutaminsäure (E81Q) die Strukturänderung kaum beeinträchtigt,<sup>139, 191</sup> scheint die geladene Seitenkette über eine intermolekulare Salzbrücke mit R57 oder K62 die Dimerbildung zu unterstützen.<sup>187, 189, 191, 197, 198, 200</sup> Die simultane Neutralisierung der Ladungen von E81 und des benachbarten D39 reduziert die Ladung des negativen Pols um 25 %, weshalb die Bindungsenergie des Dimers deutlich sinkt und der monomere Zustand stabilisiert wird.<sup>187, 202</sup> Im Gegensatz dazu wird  $\text{NRN}_{E.a.}$  durch die gemeinsame Protonierung von E81 mit E76 und E114 auf die dimere Konformation II beschränkt.<sup>191</sup> Somit scheint das Zusammenspiel von E81 mit den restlichen konservierten, sauren Aminosäuren die pH-abhängige Dimerisierung zu ermöglichen.<sup>191, 200</sup>

Die Ausrichtung von E114 variiert zwischen den gelösten Strukturen der homologen  $\text{NRN}_{E.a.}$  und  $\text{NRN}_{N.c.}$  geringfügig.<sup>200</sup> Dennoch werden in den beiden leicht unterschiedlichen Dimerisierungsschnittstellen gleichermaßen Wassermoleküle von dem geladenen Carboxylat der E114 gebunden, wodurch das Dimer bei neutralem pH-Wert destabilisiert wird.<sup>197</sup> Der  $\text{pK}_s$ -Wert von E114 (4,4) wird durch die periphere Platzierung abseits des sauren *Clusters* (Abstand ca.  $\sim 20 \text{ \AA}$ ) im Monomer kaum durch benachbarte Ladungen beeinflusst (Tabelle 2),<sup>198</sup> jedoch durch die elektrostatischen Wechselwirkungen im vorausgerichteten Dimer auf  $\text{pK}_s = 6,1 - 6,7$  angehoben.<sup>197</sup> Die resultierende Protonierung der Carboxylatgruppe von E114 neutralisiert einen Teil des Dimerisierungsinterfaces.<sup>197</sup> Trotz der geringeren Gesamtladung entsteht jedoch eine klarere dipolare Ladungsverteilung und damit korrelierend eine stärkere Dipol-Dipol Wechselwirkung.<sup>198</sup> Simultan zur Neutralisierung von E114 rotiert die Seitenkette von E114 aus dem Interface, sodass weniger

Wassermoleküle in der hydrophober werdenden Schnittstelle gebunden werden und die Dimerbindung um ca.  $3,5 \text{ kcal mol}^{-1}$  stabilisiert wird.<sup>189, 197</sup> Damit im Einklang favorisiert E114Q als einzige Einfachmutante von  $\text{NRN}_{E.a.}$  bei neutralem pH Wert den dimeren Zustand<sup>191, 197, 198</sup> und senkt bei pH 6,0 die Dissoziationsrate  $k_{\text{off}}$  des Wildtyps ( $0,30 \pm 0,05 \text{ s}^{-1}$ ) um das Fünffache auf  $k_{\text{off}}(\text{E114Q}) = 0,06 \pm 0,01 \text{ s}^{-1}$  ab.<sup>189</sup>

Neben den Asparagin- und Glutaminsäuren rückte ein konserviertes Ala an Position 69 in den Fokus, da es in der Dimerisierungsschnittstelle mit der gleichen Aminosäure der zweiten Untereinheit im direkten Kontakt ( $\sim 3 \text{ \AA}$ , PDB ID: 2LTH) steht.<sup>185</sup> Der Austausch des A69 mit einem Arginin beschränkt die Faltung der Variante A69R auf Grund der eingeführten Ladung und der sterischen Hinderung des Arginins auf die pH-unabhängige Konformation I, was als Verschiebung des Monomer-Dimer Gleichgewichts in Richtung der Monomere interpretiert wurde.<sup>185</sup>



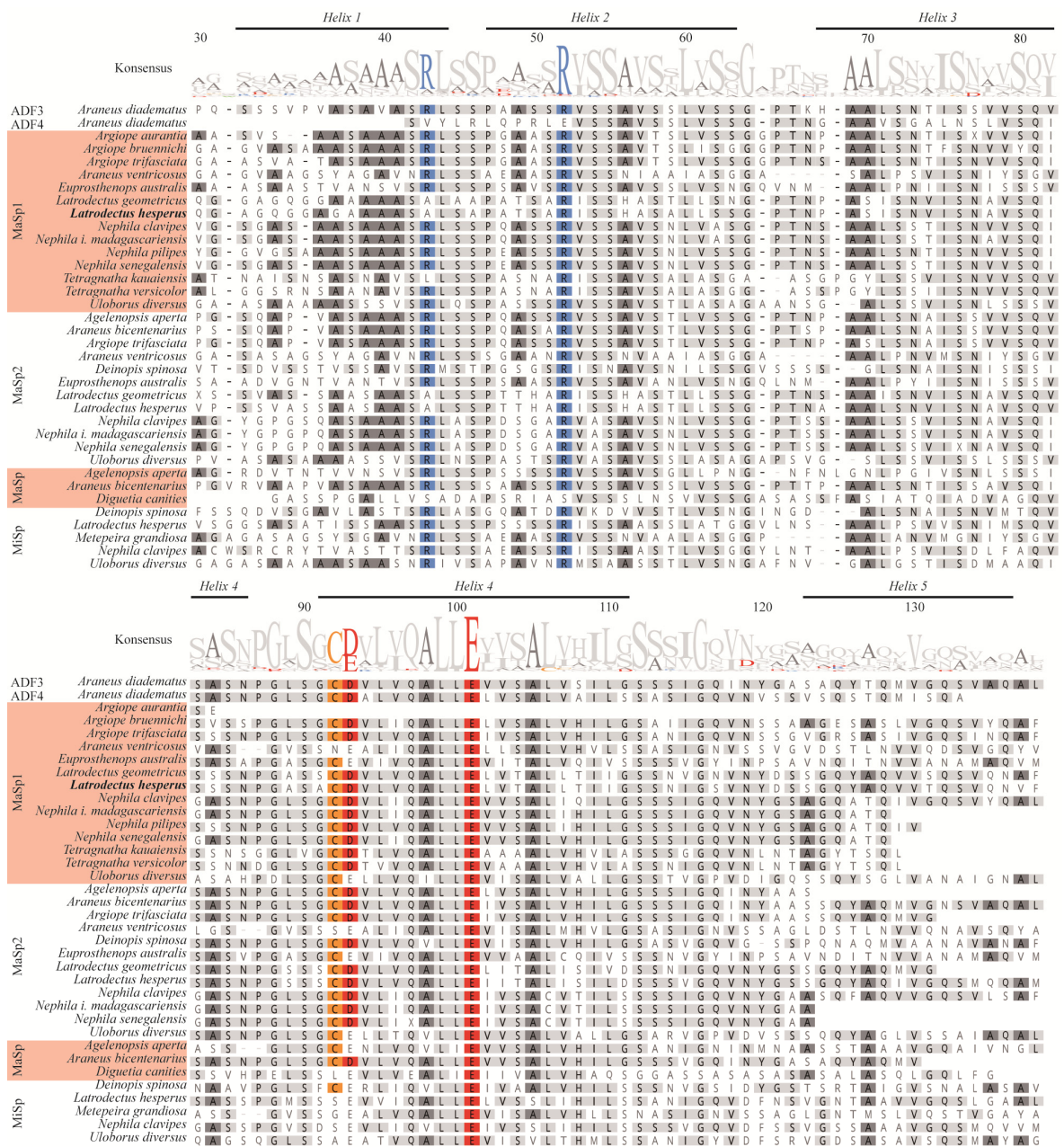
**Abbildung 8: Vorschlag eines Dimerisierungsmechanismus von  $\text{NRN}_{E.a.}$ .** Ausgehend von der Konformation I richten sich die monomeren Untereinheiten zunächst über elektrostatische Wechselwirkungen antiparallel aus. Ein sinkender pH-Wert stabilisiert über die simultane Protonierung der Seitenketten von E79 und E114 die Konformation II, in der W9 exponiert und ein schwach assoziiertes Dimer gebildet wird. Weitere Ansäuerung begünstigt die Protonierung von E81, woraufhin die geringfügig veränderte Konformation II\* die Dimerbindung verstärkt (Abbildung angelehnt an Kronqvist *et al.*<sup>191</sup>).

Anhand der größtenteils auf strukturellen Untersuchungen basierenden Erkenntnisse wurde von Kronqvist *et al.* ein Dimerisierungsmodell für  $\text{NRN}_{E.a.}$  vorgeschlagen (Abbildung 8).<sup>191</sup> Bei neutralem pH-Wert und in der Anwesenheit von NaCl liegt  $\text{NRN}_{E.a.}$  als stabiles Monomer in Konformation I vor, in der das W9 im hydrophoben Inneren des Fünf-Helix-Bündels verborgen ist. Eine sinkende NaCl-Konzentration verstärkt die elektrostatischen Interaktionen zwischen D39/E81 und R57/K62, sodass sich die dipolaren Untereinheiten antiparallel ausrichten. Die Annäherung der Untereinheiten lässt die  $\text{pK}_s$ -Werte von E76 und E114 ansteigen<sup>197</sup> und ermöglicht zusammen mit einer leichten Ansäuerung die Protonierung der beiden Glutamatseitenketten, woraufhin die schwach assoziierten Dimere die dichter gepackte Konformation II einnehmen und das W9 an die Oberfläche der Struktur verschoben wird. Das weitere Absinken des pH-Wertes führt zur zusätzlichen Protonierung

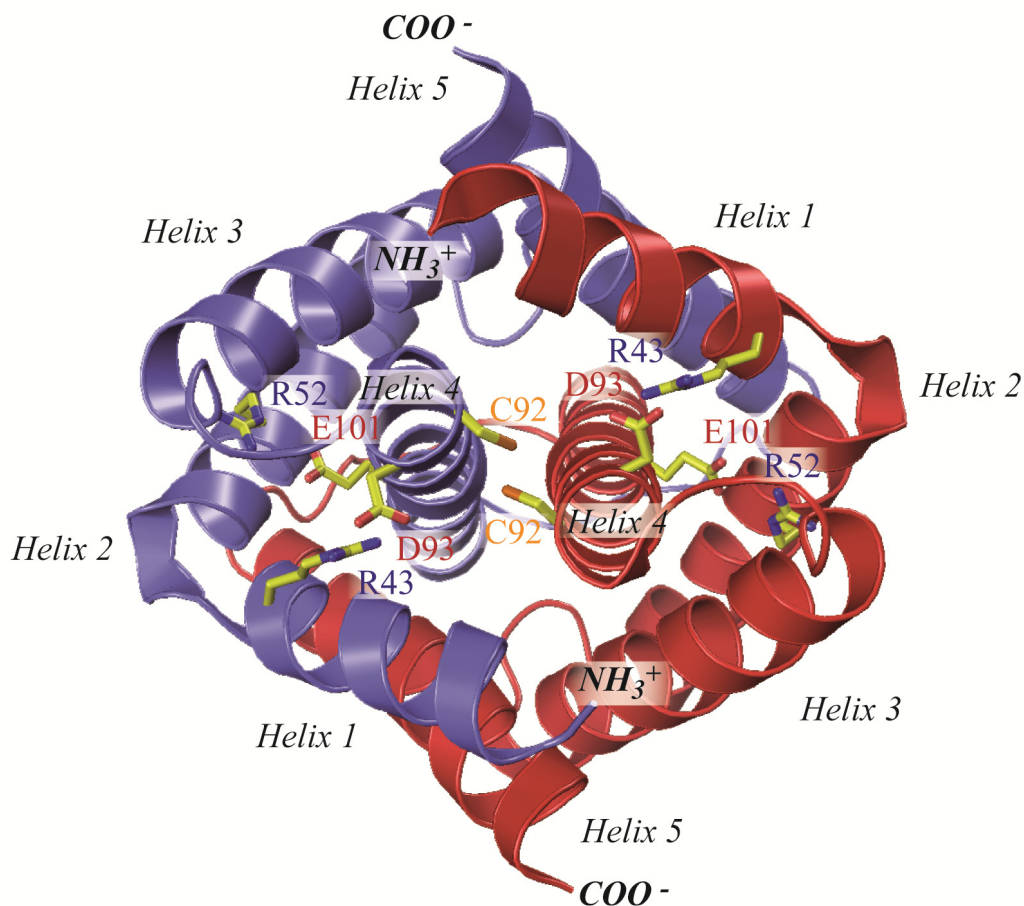
von E81 und Stabilisierung des Dimers in der leicht angepassten Konformation II\*.<sup>191</sup> Die pH- und Salz-sensitive Dimerisierung von NRN initiiert die Spidroinassembly somit in dem Protonen- und Salzgradienten, der entlang der natürlichen Spinnndrüse durchlaufen wird.<sup>120-123, 130-133</sup>

#### 1.7.5 CARBOXY-TERMINALE DOMÄNE

Der dreiteilige Aufbau der Spidroine wird durch die nicht-repetitive carboxy-terminale Domäne (NRC) vervollständigt, deren Aminosäuresequenz mit ~75 % Identität<sup>203</sup> zwischen verschiedenen MaSp und MiSp einen annähernd so hohen Konservierungsgrad aufweist wie NRN.<sup>39, 98, 135, 140, 203-206</sup> Verglichen mit NRC von MaSp weisen die Flagelliform- und Aciniform-Spidroine eine signifikant divergente Primärstruktur (27 % Identität bei *N. clavipes*) auf, dennoch nehmen auch deren NRC eine ähnliche Faltung ein.<sup>203, 206, 207</sup> Zwischen den verschiedenen NRC aus MaSp und MiSp liegen zwei basische (R43, R52) und zwei saure Aminosäuren (D93, E101) stark konserviert und ein Cystein (C92) relativ konserviert vor (Abbildung 9), sodass diesen Resten offensichtlich eine besondere Funktion zukommt. Ähnlich wie NRN (siehe 1.7.4) nehmen die homologen NRC aus MaSp und MiSp eine Fünf-Helix-Bündel Struktur ein.<sup>138</sup> In der gelösten Struktur von NRC<sub>A.d.</sub> aus *Araneus diadematus* Fibroin 3 sind die konservierten Ladungen jedoch nicht wie in NRN Dipol-ähnlich auf der Oberfläche exponiert, sondern stabilisieren das Protein im Inneren über zwei intramolekulare Salzbrücken R43–D93 und R52–E101 (Abbildung 10).<sup>138</sup> In der Struktur bilden zwei NRC Untereinheiten über die stark hydrophobe Helix 4 (Abbildung 9) ein parallel orientiertes Dimer aus, das in den meisten MaSp über eine intermolekulare Disulfidbrücke (C92–C92, Distanz 2 Å) kovalent gebunden wird (Abbildung 10).<sup>127, 138, 203, 205, 208</sup> Zusätzlich wird das Dimer über nichtkovalente Wechselwirkungen zwischen Helix 1 der ersten und Helix 5 der zweiten Untereinheit in einer Klammer-ähnlichen Ausrichtung fixiert (Abbildung 10). Interessanterweise übertrifft die Hydrophobizität der NRC-Sequenz sogar die der repetitiven Kerndomäne,<sup>127</sup> wobei die hydrophoben Aminosäuren (u.a. Valin, Leucin, Isoleucin) im gefalteten Zustand durch die hydrophilen Aminosäuren (u.a. Glutamin, Serin) abgeschirmt sind.<sup>138</sup>



**Abbildung 9: Sequenzalignment der C-terminalen Domänen verschiedener Spidroine.** Die Aminosäuresequenz der C-terminalen Domänen 36 verschiedener Spidroine wurde mittels Geneious abgeglichen.<sup>183</sup> Aminosäuren wurden ab einer Identität von 25 % in der Konsensussequenz berücksichtigt (hellgrau hinterlegt, Alanin dunkelgrau, Cystein orange, saure Aminosäuren rot und basische blau). Die Größe der Aminosäurebuchstaben in der Konsensussequenz entspricht der prozentualen Identität der Aminosäure. Mit Hilfe von Jpred 4 wurden für die Konsensussequenz die helikalen Bereiche berechnet.<sup>184</sup> ADF, *Araneus diadematus* Fibroin 3; MaSp1, Major Ampullate Spidroin 1; MaSp2, Major Ampullate Spidroin 2; MiSp, Minor Ampullate Spidroin.



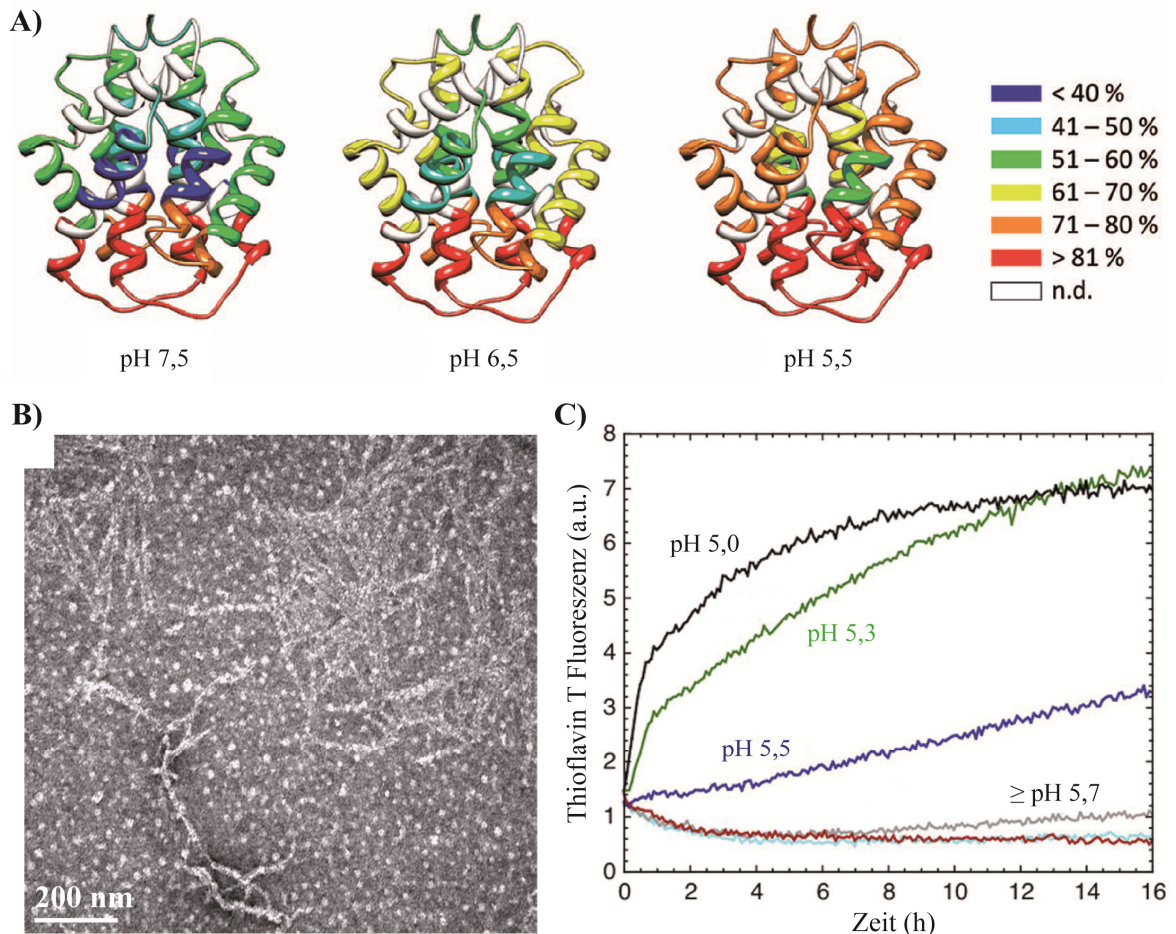
**Abbildung 10: NMR-Struktur von NRC<sub>A.d.</sub>** Die Faltung von NRC<sub>A.d.</sub> aus *Araneus diadematus* Fibroin 3 (PDB ID: 2KHM)<sup>138</sup> zeigt beispielhaft die konservierte Fünf-Helix-Bündel Struktur, die über zwei intramolekulare Salzbrücken (R43–D93 und R52–E101) stabilisiert wird. Zwei NRC-Untereinheiten (blau und rot) sind über Helix 4 zu einem parallelen Dimer angeordnet, das über die intermolekulare Disulfidbrücke C92–C92 kovalent verknüpft ist. Die Abbildung wurde mit PyMOL erstellt.<sup>141</sup> Farbmarkierung der Atome: C, gelb; N, blau; O, rot; S, orange.

Das hydrierte NRC bildet zusammen mit dem ebenfalls hydrophilen NRN die lösliche Schale der Mizellen-ähnlichen Strukturen aus, um eine verfrühte Aggregation während der Lagerung der hochkonzentrierten Spidroinlösung in der Spinnampulla zu verhindern.<sup>124–128</sup> Während die Spidroine den S-förmigen und konisch zulaufenden Spinntrakt passieren, induzieren die auftretenden Scherkräfte die Exponierung hydrophober Bereiche aus dem Inneren von NRC.<sup>138</sup> Dieser Prozess richtet die Sekundärstruktur motive der Spidroine während der Assemblierung aus und forciert die Koaleszenz der Mizellen.<sup>124, 125, 129</sup> Somit besitzt NRC sowohl bei der Lagerung als auch der Assemblierung der Spidroine eine enorm wichtige Funktion.<sup>107, 127, 138, 203</sup>

Bei einem extrem niedrigen und nicht-physiologischen pH-Wert von < 5,5 falten die isolierten NRC<sub>A.d.</sub>,<sup>138</sup> NRC<sub>A.v.</sub> (NRC aus *A. ventricosus* MiSp),<sup>131</sup> und NRC<sub>N.c.</sub> (NRC aus *N. clavipes* MaSp1)<sup>209</sup> in eine destabilisierte und flexiblere *molten globule* Struktur (englisch,



Intermediat der Proteinfaltung), die u.a. durch den verstärkten Deuteriumaustausch bei pH 5,5 mittels NMR Spektroskopie nachgewiesen wurde (Abbildung 11A).<sup>131</sup> Die hydrophoben Bereiche von NRC<sub>A.v.</sub> sind in dem *molten globule* Zustand leichter zugänglich, sodass sich nach mehrstündiger Inkubation bei  $\leq$  pH 5,5 Fibrillen-ähnliche Aggregate mit erhöhtem  $\beta$ -Faltblattanteil ausbilden (Abbildung 11B).<sup>131</sup> Dabei indiziert die erhöhte Affinität zu Thioflavin-T (Abbildung 11C) und Kongorot die Ausbildung von Amyloid-ähnlichen Strukturen.<sup>131, 209</sup>



**Abbildung 11: NRC<sub>A.v.</sub> faltet bei pH < 5,5 in *molten globules*, aus denen Amyloid-ähnliche Aggregate entstehen.** A) Der Deuteriumaustausch in NRC<sub>A.v.</sub> wurde pH-abhängig untersucht. Der Deuteriumanteil von NRC<sub>A.v.</sub> bei pH 7,5, pH 6,5 und pH 5,5 wird durch das Farbschema codiert. B) Das Transmissionselektronenmikroskopiebild zeigt, dass NRC<sub>A.v.</sub> nach Inkubation bei pH 5,5 Amyloid-ähnliche Fasern bildet. C) Die ansteigende Fluoreszenzintensität von Thioflavin T bei  $\leq$  pH 5,5 signalisiert die verstärkte Faltung von NRC<sub>A.v.</sub> in  $\beta$ -Faltblattstrukturen. Modifiziert nach Andersson *et al.*<sup>131</sup>

#### 1.7.6 WECHSELWIRKUNGEN

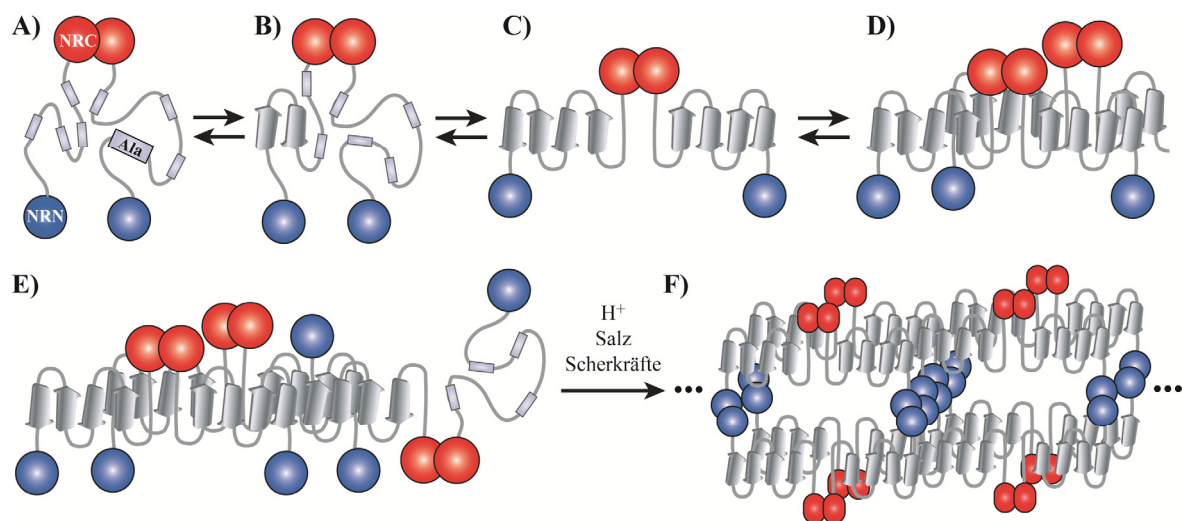
#### ZWISCHEN

#### REKOMBINANTEN

#### ASSEMBLIERUNGSDOMÄNEN

Die Kerndomäne der rekombinanten Spinnenseidenproteine (eADF3 und eADF4) weist im löslichen Zustand keine definierte Tertiärstruktur auf<sup>210</sup> und wird somit ähnlich zu den natürlichen Spidroinen als intrinsisch ungefaltet angesehen.<sup>211, 212</sup> Die Assemblierung von

eADF wird durch Zugabe kosmotroper Kalium- und Phosphationen ausgelöst.<sup>210, 213</sup> Im Vergleich zu globulär gefalteten Strukturen wie den NRN- und NRC-Domänen, erleichtert die flexible Konformation der Kerndomäne den Übergang von einer löslichen Struktur in unlösliche,  $\beta$ -Faltblatt-reiche Nanofibrillen.<sup>212, 214</sup> Die Assemblierung von eADF4 beginnt mit einer Nukleationsphase, in der ein größtenteils unstrukturiertes Monomer (Abbildung 12A) in den miteinander verbunden Modulen antiparallele  $\beta$ -Faltblätter ausbildet (Abbildung 12B, C).<sup>210</sup> Anschließend interagieren die Proteine über hydrophobe Wechselwirkungen der exponierten Alaninseitenketten und lassen einen Assemblierungsnukleus entstehen (Abbildung 12D).<sup>215</sup>



**Abbildung 12: Model zur Assemblierung der rekombinanten eADF Spidroine.** A) Die intrinsisch ungefaltete Kerndomäne enthält sich wiederholende Poly-Ala Bereiche und wird von globulär gefalteten NRN und NRC-Domänen flankiert. B-C) Benachbarte Ala-Bereiche bilden  $\beta$ -Faltblätter aus und (D) bilden über hydrophobe Wechselwirkungen mit weiteren  $\beta$ -Faltblatt-reichen Spidroinen den Nukleus aus. E-F) Durch die Anlagerung weiterer Spidroine wachsen die  $\beta$ -Faltblatt-reichen Bereiche der Fibrille an (angelehnt an Humenik *et al.*).<sup>210</sup> Sich ändernde chemische und mechanische Bedingungen beeinflussen die Tertiärstruktur der beiden terminalen Domänen und die Quartärstruktur von NRN, was die Assemblierung der Spidroine unterstützt.

In der darauffolgenden Elongationsphase werden dem Nukleus bzw. dem daraus hervorgehenden Fibrillenende weitere Spidroine hinzugefügt (Abbildung 12E). Der Nukleus bzw. die Fibrille kann aktiv die Struktur der Monomere in Lösung ändern, die damit selbst zum Nukleus werden (Abbildung 12F). Die von eADF(C16) gebildeten Nanofibrillen besitzen Cross- $\beta$ -Strukturen, da sie ein typisches Röntgenbeugungssignal und eine hohe Affinität zu Kongorot und Thioflavin T aufweisen.<sup>216</sup>

Der Assemblierungsprozess der rekombinanten Spidroine hängt neben der Art und Anzahl der repetitiven Module<sup>210</sup> auch signifikant von der Präsenz der flankierenden terminalen

Domänen ab. Da das Molekulargewicht der Kerndomäne bzw. die Anzahl der Module direkt mit der Kinetik der Spidroinassembly korreliert,<sup>210</sup> unterstützt die NRC-vermittelte Dimerisierung der Spidroine auf Grund der Verdoppelung des Molekulargewichts die Assembly. Zudem werden durch die parallele Ausrichtung des NRC Dimers die repetitiven Module zweier Spidroine bereits in unmittelbare Nähe gebracht.

Während das von Insektenzellen exprimierte eADF4 ohne terminale Domäne unspezifische Aggregate ausbildet, vermittelt die Fusion von NRC<sub>A.d.</sub> die Assembly von eADF4 in Nanofibrillen mit kristallinen Bereichen.<sup>153</sup> Die Scherkraft- und Phosphat-induzierte Destabilisierung der NRC<sub>A.d.</sub> Faltung bringt hydrophobe Bereiche aus dem Domäneninneren an die Oberfläche und treibt die auf hydrophoben Wechselwirkungen basierende Assembly der Kerndomäne an.<sup>138</sup> Rheologische Messungen belegen, dass rekombinante Spidroine eine NRC Domäne enthalten müssen, um einen Scherkraft-induzierten Anstieg der Viskosität zu erfahren.<sup>217, 218</sup> Die NRC Domäne beschleunigt *in vitro* die Assembly und unterstützt die parallele Ausrichtung der rekombinanten Kerndomänen in Fibrillenbündel mit entlang der Faserachse angeordneten  $\beta$ -Faltblättern.<sup>138, 207, 219, 220</sup> Während NRC die Entstehung geordneter Seidenfäden ermöglicht, steuert NRN die Spidroinassembly in einer pH- und Salz-abhängigen Art.<sup>219</sup> Da über die parallelen NRC-Dimere bereits zwei Spidroine miteinander verbunden sind,<sup>138</sup> entstehen durch die antiparallele Dimerisierung von NRN lange Spidroinketten und quervernetzte Mikrokristalle (Abbildung 12).<sup>140</sup>

Zudem wurde spekuliert, ob die durch starke Ansäuerung verursachte *molten globule* Struktur des isolierten NRC ebenfalls als  $\beta$ -Faltblatt-reicher Nukleus fungieren könnte, um die Assembly der Kerndomäne anzutreiben.<sup>131, 209</sup> Gegen diese Hypothese spricht, dass die Bildung der Amyloid-Strukturen einen nicht-physiologischen pH-Wert von  $\leq$  pH 5,0 benötigt, eine signifikant langsamere Kinetik (h) als die Spidroinassembly (ms) aufweist und der Heteronukleus aus zwei Domänen zu einer unspezifischeren Assembly führen könnte.<sup>131, 209</sup>



## 2. ZIELSETZUNG

Das Ziel dieser Arbeit war, die während der Spidroinassembly ablaufenden Mechanismen auf molekularer Ebene zu untersuchen. In vorherigen Arbeiten wurde demonstriert, dass die nicht-repetitive N-terminale Domäne (NRN) über die pH- und NaCl-abhängige, antiparallele Dimerisierung und Konformationsänderung ihrer Fünf-Helix-Bündel Struktur die Assemblierung der Spinnenseidenproteine in der Spinndrüse unterstützt.<sup>139</sup> Der molekulare Mechanismus der pH- und Salz-induzierten Strukturänderung und Dimerbildung sowie der Zusammenhang zwischen beiden Prozessen waren zu Beginn dieser Arbeit jedoch nicht bekannt. Um die Protonierung hochkonservierter Asparagin- (D39) und Glutaminsäuren (E76, E81, E114) zu imitieren, sollten in dieser Arbeit mit Hilfe ortsspezifischer Mutagenese Varianten von NRN<sub>1<sub>L.h.</sub></sub> aus *Latrodectus hesperus* MaSp1 erstellt werden. Dabei sollten einzelne oder multiple negative Ladungen der Asparagin- und Glutaminsäuren durch Austausch mit deren nicht-titrierbaren Analoga Asparagin und Glutamin neutralisiert oder durch Substitution mit Arginin umgekehrt werden. Die rekombinanten NRN-Varianten sollten biotechnologisch in dem Wirtsorganismus *E. coli* hergestellt und mit Hilfe einer zu optimierenden, chromatographischen Reinigungsstrategie von Verunreinigungen befreit werden. Die Struktur der Varianten von NRN<sub>1<sub>L.h.</sub></sub> sollte mit Hilfe verschiedener spektroskopischer Methoden (u.a. Zirkulardichroismus, Fluoreszenz) bei unterschiedlichen pH-Werten und Salzkonzentrationen untersucht werden. Zudem sollte in Kooperation mit dem Lehrstuhl für Biopolymere (Prof. Dr. Schwarzing, Prof. Dr. Rösch, Universität Bayreuth), mittels Kernspinresonanz-Spektroskopie die Struktur einer Dreifachmutante (D39N-E76Q-E81Q) mit neutralisiertem sauren Cluster bestimmt werden, um den Einfluss der akkumulierten Ladungen auf die Faltung aufzudecken. Da in vorherigen Veröffentlichungen das Monomer-Dimer-Gleichgewicht nicht direkt gemessen, sondern indirekt über die Konformationsänderung beurteilt wurde,<sup>180, 191, 200</sup> sollten zudem die Molekulargewichte der Varianten über Lichtstreuungsexperimente und der Zusammenhang (Voraussetzung, Folge, Unabhängigkeit) zwischen Strukturänderung und Dimerbildung bestimmt werden.

Im nächsten Schritt sollte anhand von Fusionsproteinen aus NRN und einer Kerndomäne das Wechselspiel zwischen den beiden Domänen charakterisiert werden. Zum einen sollten die Folgen der unstrukturierten Kerndomäne auf die Dimerisierung von NRN analysiert und zum anderen der Einfluss des Monomer-Dimer Gleichgewichts von NRN auf die Assemblierung der Kerndomäne untersucht werden. Dafür sollten Monomer- und Dimer-stabilisierte Varianten von NRN<sub>1<sub>L.h.</sub></sub> mit einer rekombinanten Kerndomäne (AQ), die nach

der Konsensussequenz der Kerndomäne von *Araneus diadematus* Fibroin 3 entworfen wurde, fusioniert werden. Neben der Faltung und dem Monomer-Dimer-Gleichgewichts der Fusionsproteine sollte insbesondere deren pH, Phosphat und Scher-induzierte Assemblierung charakterisiert werden.

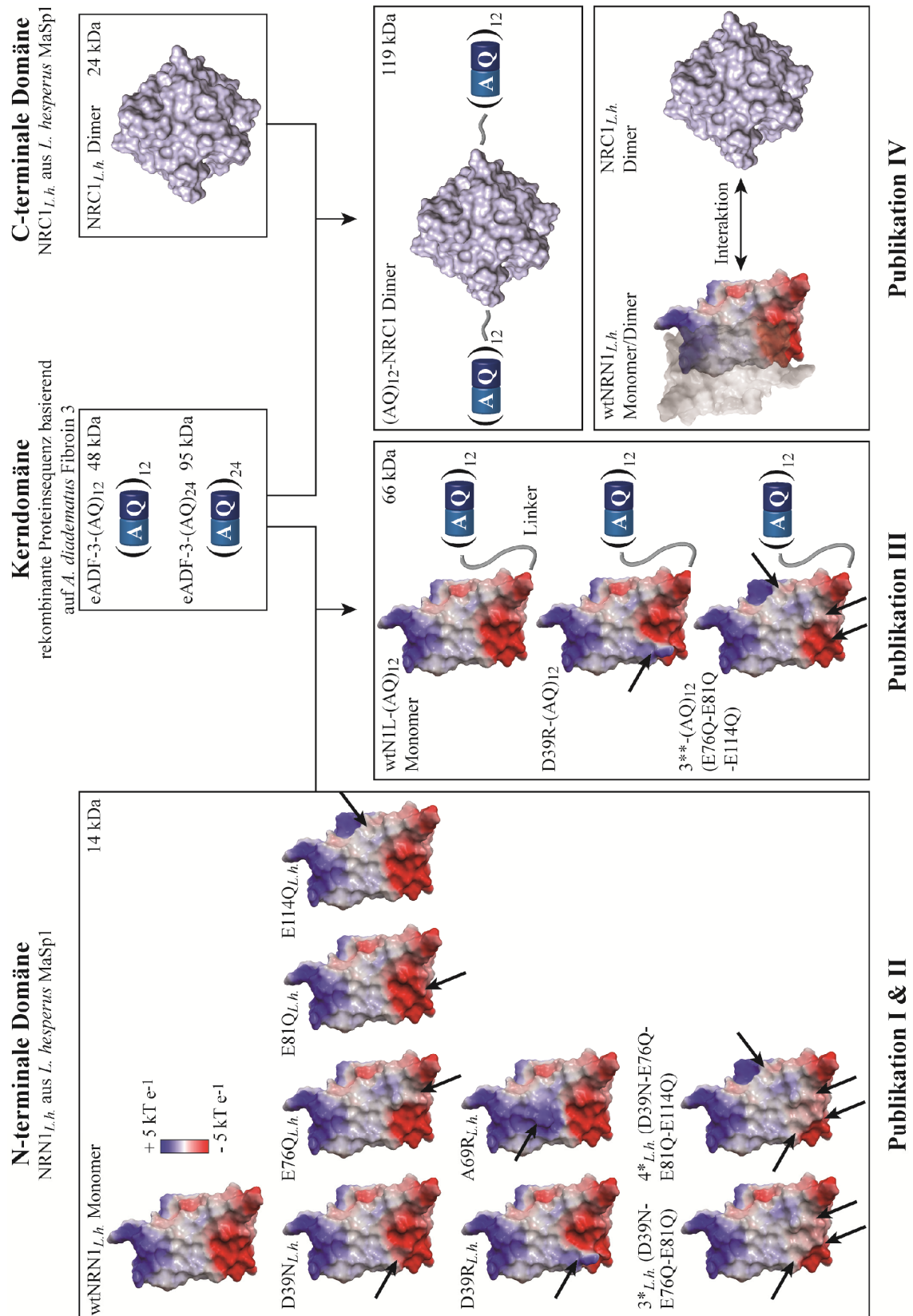
Es wird angenommen, dass NRN zusammen mit der nicht-repetitiven C-terminalen Domäne (NRC) präferenziell die hydrophile Oberfläche der Mizellen-ähnlichen Strukturen während der Lagerung der Spidroine im Reservoir der Spinndrüse ausbildet. Da die molekularen Wechselwirkungen zwischen den beiden terminalen Domänen nicht bekannt sind, sollten im Rahmen dieser Arbeit die intermolekularen Wechselwirkungen zwischen NRN<sub>L.h.</sub> und NRC<sub>L.h.</sub> aus *L. hesperus* MaSp1 mit diversen Analysetechniken untersucht werden. Für NRC wurde zudem die Ausbildung Amyloid-ähnlicher Fasern bei pH ≤ 5 berichtet,<sup>131, 209</sup> die gegebenenfalls als  $\beta$ -Faltblatt-reiche Nuklei die Assemblierung der hydrophoben Bereiche der Kerndomäne beschleunigen könnten. Um diese Hypothese zu überprüfen, sollte die Struktur des isolierten NRC<sub>L.h.</sub> sowie die Assemblierung von Fusionskonstrukten aus NRC<sub>L.h.</sub> und der rekombinanten (AQ)-Kerndomäne pH-abhängig charakterisiert werden.

### 3. SYNOPSIS

Diese Arbeit umfasst die in vier Publikationen (siehe 5) beschriebene Charakterisierung rekombinanter Spidroine, die aus einzelnen und multiplen Assemblierungsdomänen bestehen (Abbildung 13). Sämtliche Spidroine wurden erfolgreich mit Hilfe des Wirtsorganismus *E. coli* hergestellt und über selektive Reinigungsstrategien, die spezifisch für jedes Konstrukt entwickelt und optimiert wurden, von Verunreinigungen separiert.

#### 3.1 ÜBERSICHT ÜBER UNTERSUCHTE ASSEMBLIERUNGSDOMÄNEN VON SPINNENSEIDENPROTEINEN

In Publikation I wurden Varianten der NRN<sub>L.h.</sub>-Domäne aus *Latrodectus hesperus* MaSp1 (siehe Hagn *et al.*)<sup>140</sup> erstellt, in denen einzelne (D39N<sub>L.h.</sub>, E76Q<sub>L.h.</sub>, E81Q<sub>L.h.</sub>, E114Q<sub>L.h.</sub>) oder mehrere (3\*<sub>L.h.</sub> = D39N-E76Q-E81Q, 4\*<sub>L.h.</sub> = D39N-E76Q-E81Q-E114Q) der hochkonservierten Asparagin- bzw. Glutaminsäuren durch die nicht-titrierbaren Analoga Asparagin und Glutamin ersetzt wurden (Abbildung 13).<sup>221</sup> Die Struktur von Variante 3\*<sub>L.h.</sub> wurde mit Hilfe von Kernspinresonanz- (NMR) Spektroskopie gelöst und deren Resonanzzuordnung in Publikation II beschrieben.<sup>222</sup> Neben der Neutralisierung der Carboxylgruppen wurden auch eine Ladung des sauren *Clusters* umgekehrt (D39R<sub>L.h.</sub>) bzw. eine Ladung in das Dimerisierungsinterface (A69R<sub>L.h.</sub>) eingeführt (Abbildung 13).



**Abbildung 13: Übersicht der in dieser Arbeit charakterisierten rekombinanten Spidroine, die aus einer oder zwei Assemblierungsdomänen bestehen.** In Publikation I wurden mit Hilfe ortsspezifischer Mutagenese Varianten der N-terminalen Domäne (wtNRN1<sub>L.h.</sub>) aus *Latrodectus hesperus* MaSp1 erstellt, in denen einzelne oder multiple konservierte Aminosäuren des Dimerisierungsinterfaces ausgetauscht und das elektrostatische Oberflächenpotential modifiziert wurde (markiert mit Pfeil).<sup>221</sup> In Publikation II wurde mittels Kernspinresonanz-Spektroskopie die Struktur der Dreifachmutante 3\*<sub>L.h.</sub> (D39N-E76Q-E81Q) mit neutralisiertem sauren Cluster bestimmt.<sup>222</sup> Anschließend wurde in Publikation III das Zusammenspiel einer rekombinanten Kerndomäne (AQ)<sub>12</sub> mit im monomeren bzw. dimeren Zustand stabilisierten Varianten von NRN1<sub>L.h.</sub> untersucht.<sup>223</sup> Die C-terminale Assemblierungsdomäne NRC1<sub>L.h.</sub> aus *L. hesperus* MaSp1 wurde in Publikation IV einzeln und als Fusionsprotein mit der rekombinanten Kerndomäne (AQ)<sub>12</sub> analysiert.<sup>224</sup> Zudem wurde die Interaktion zwischen NRC1<sub>L.h.</sub> und monomeren sowie dimeren NRN1<sub>L.h.</sub> charakterisiert. Das elektrostatische Oberflächenpotential der N- (PDB ID: 2N3E)<sup>222</sup> und C-terminalen Domäne (PDB ID: 2KHM)<sup>138</sup> wurde mit Hilfe von PyMOL dargestellt und modifiziert.<sup>141</sup> Die Struktur von NRC1<sub>L.h.</sub> wurde bisher nicht bestimmt und wurde in dieser Abbildung durch die Struktur des homologen NRC<sub>A.d.</sub> aus ADF3 dargestellt. Schema enthält modifizierte Teilabbildung aus Bauer *et al.*<sup>223</sup> mit Copyright © 2017 American Chemical Society.

Durch Analyse der Sekundär-, Tertiär- und Quartärstruktur wurden im monomeren (D39R<sub>L.h.</sub>) und dimeren Zustand (4\*<sub>L.h.</sub>) stabilisierte Varianten von wtNRN1<sub>L.h.</sub> identifiziert und in Publikation III mit einer rekombinanten Kerndomäne fusioniert,<sup>223</sup> um den Einfluss der Multimerblöcke auf die Dimerisierung von NRN1<sub>L.h.</sub> zu analysieren. Da in der Zwischenzeit von da Silva *et al.* aufgedeckt wurde, dass die Protonierung von D39 die Bindungsenergie des Dimers nicht beeinflusst und die Asparaginsäure im Dimer eine starke Tendenz zur Abstoßung des Protons aufweist,<sup>198</sup> wurde die in 4\*<sub>L.h.</sub> (D39N-E76Q-E81Q-E114Q) enthaltene Mutation D39N exkludiert und die Variante 3\*\*<sub>L.h.</sub> (E76Q-E81Q-E114Q) erstellt. Die rekombinante Kerndomäne wurde in vorhergehenden Arbeiten aus der Konsensussequenz der repetitiven Poly-Ala (Modul A) und GPGQQ Blöcke (Modul Q) von Major Ampullate Fibroin 3 (ADF3) aus *Araneus diadematus* MaSp2 abgeleitet<sup>39</sup> und nach Anpassung an den Wirtsorganismus modular zu dem *engineered* (englisch, gentechnisch verändert) eADF3-(AQ)<sub>n</sub> multimerisiert.<sup>160</sup> Die bisher nicht identifizierte Aminosäuresequenz von NRN<sub>A.d.</sub> aus ADF3 (NCBI GenBank, Stand 31.03.2017, Zugangscode ADF3: AAC47010, ADF4: AAC47011)<sup>39, 225</sup> wurde durch die Wildtypsequenz und modifizierten Sequenzen des homologen NRN1<sub>L.h.</sub> ersetzt und über eine kurze Linkerregion (L) mit zwölf Wiederholungen des rekombinanten (AQ)-Blockes zu Varianten von N1L-(AQ)<sub>12</sub> zusammengefügt.

Der vierte Teilbereich dieser Arbeit befasst sich mit der individuellen, nicht-repetitiven C-terminalen Domäne NRC1<sub>L.h.</sub> aus *Latrodectus hesperus* MaSp1.<sup>224</sup> Zudem wurden die intermolekularen Wechselwirkungen zwischen den individuellen Assemblierungsdomänen

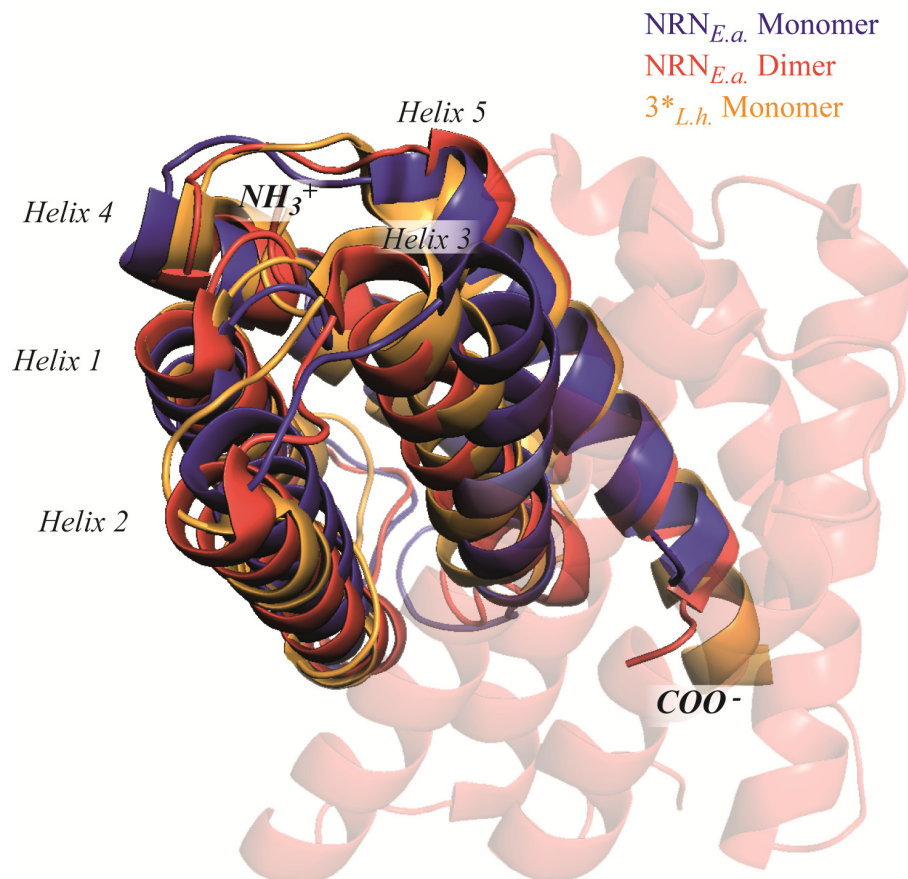
NRC1<sub>L.h.</sub> und NRN1<sub>L.h.</sub> analysiert. Dabei wurde überprüft, ob die beiden konservierten molekularen Schalter die Spidroinassembly im Zusammenspiel oder unabhängig voneinander steuern. Um die zuvor für homologe NRC beschriebene<sup>131, 209</sup> Ausbildung Amyloid-ähnlicher und  $\beta$ -Faltblatt-reicher Fasern zu verifizieren, wurde zudem die Struktur und Stabilität von NRC1<sub>L.h.</sub> im stark sauren, nicht-physiologischen pH-Bereich charakterisiert. Der Zusammenhang zwischen dem pH-abhängigen Verhalten von NRC1<sub>L.h.</sub> und der Assemblierung der Kerndomäne wurde mit Hilfe des Fusionskonstrukts (AQ)<sub>12</sub>-NRC1 bestimmt.

### 3.2 STRUKTURELLE UND FUNKTIONELLE CHARAKTERISIERUNG DER NICHT-REPETITIVEN N-TERMINALEN DOMÄNE NRN

Im natürlichen Spinnprozess wird die Assemblierung löslicher MaSp zu einem stabilen Faden unter anderem über die pH- und Salz-abhängige Dimerisierung und Strukturänderung von NRN gesteuert. Die Tertiärstruktur von wtNRN1<sub>L.h.</sub> veränderte sich zwischen pH 7,2 und pH 6,0, sodass NMR-spektroskopische Untersuchungen stark unterschiedliche <sup>1</sup>H-<sup>15</sup>N-HSQC (englisch, *heteronuclear single quantum coherence*) Spektren ergaben. Ausgehend von einer monomeren Konformation I, die durch einen neutralen pH-Wert stabilisiert wird, reorganisierte sich das Helix-Bündel nach Abfallen des pH-Wertes auf  $\leq 6,0$  in die stabilere, dimere Konformation II. Dabei wurde das einzige Tryptophan an Position 9 (W9) in eine hydrophilere Umgebung verschoben, was sich in einer reduzierten Absorption des zirkularpolarisierten Lichts im Nah-UV Bereich und einem bathochrom (rot-) verschobenen Fluoreszenzmaximum ausdrückte. Die pH-induzierte Konformationsänderung wurde sowohl durch die selektive Ladungsumkehr im Dimerisierungsinterface (D39R<sub>L.h.</sub>) als auch den Austausch eines Alanins, das in der Dimerisierungsschnittstelle mit der gleichen Aminosäure der zweiten Untereinheit in direktem Kontakt steht, mit einem positiv geladenen Arginin (A69R<sub>L.h.</sub>) nahezu vollständig unterdrückt. Die pH-unabhängige Restriktion in Konformation I wurde übereinstimmenderweise für die homologen Varianten aus NRN<sub>E.a.</sub> (D40K<sub>E.a.</sub> und A72R<sub>E.a.</sub>) berichtet.<sup>185, 191</sup>

Da die Neutralisierung der kumulierten Ladungen D39, E76 und E81 (3\*<sub>L.h.</sub>) den strukturellen Übergang in einen saureren pH-Bereich verschob, scheinen die pK<sub>S</sub>-Werte von D (4,0) und E (4,4)<sup>192, 193</sup> durch die Zusammenlagerung auf den pH-Bereich des Spinntrakts angehoben zu werden.<sup>197, 198</sup> Um die Funktion des sauren *Clusters* auf die pH-induzierte Strukturänderung zu verifizieren, wurde in Publikation II in Zusammenarbeit mit dem Lehrstuhl für Biopolymere (Prof. Dr. Schwarzing, Prof. Dr. Rösch, Universität Bayreuth)

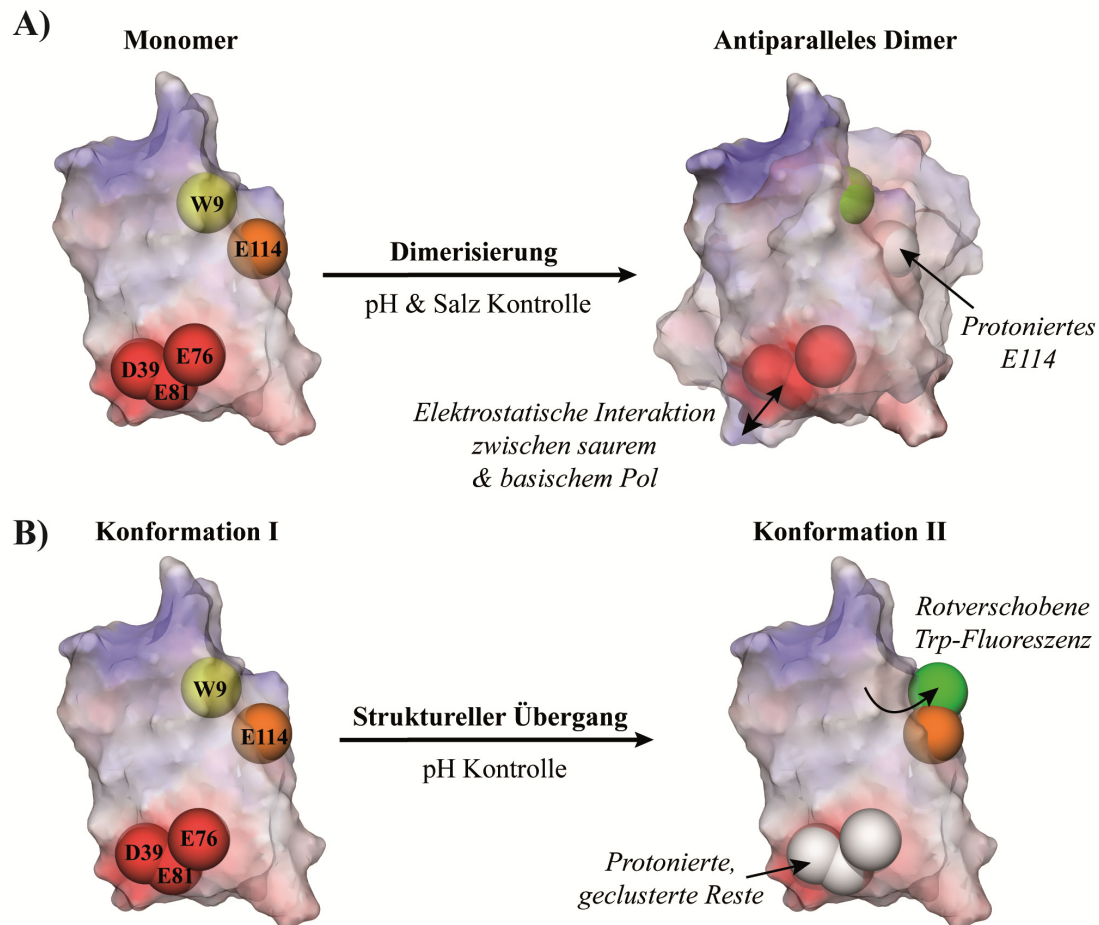
die Struktur von 3\*<sub>L.h.</sub> mittels NMR-Spektroskopie ermittelt (PDB ID: 2N3E).<sup>222</sup> Ein jCE/jFATCAT<sup>226-228</sup> Strukturvergleich von 3\*<sub>L.h.</sub>, deren Struktur bei Monomerstabilisierenden Bedingungen gelöst wurde, ergab eine deutlich höhere Ähnlichkeit mit der dimeren Konformation II (rmsd 2,02 (englisch, *root mean square deviation*)) als mit der monomeren Konformation I (rmsd 2,37) (Abbildung 14). Dies bestätigte den für NRN<sub>E.a.</sub> suggerierten Dimerisierungsmechanismus, in dem E76 und E81 an der Konformationsänderung beteiligt sind (siehe Abbildung 8).<sup>139, 185, 187</sup>



**Abbildung 14: NMR spektroskopische Untersuchung der NRN<sub>L.h.</sub>-Variante 3\*<sub>L.h.</sub> (D39N-E76Q-E81Q).** Überlagerung der energieärmsten Struktur der monomeren NRN<sub>L.h.</sub>-Variante 3\*<sub>L.h.</sub> (orange, pH 7,2, PDB ID: 2N3E) mit der monomeren (blau, pH 7,2, PDB ID: 2LPJ)<sup>185</sup> und dimeren (rot, pH 5,5, PDB ID: 2LTH)<sup>191</sup> Struktur von NRN<sub>E.a.</sub>. Die Strukturen wurden mit PyMOL dargestellt.<sup>141</sup>

Um den Zusammenhang zwischen der Strukturänderung und der Dimerisierung aufzuklären, wurden die Molekulargewichte der NRN<sub>L.h.</sub> Varianten mit Hilfe von SEC-MALS (englisch, *size exclusion chromatography*, Gel-Permeations-Chromatographie; *multi-angle light scattering*, Mehrwinkel-Lichtstreuung) ermittelt. Das Monomer-Dimer Gleichgewicht von wtNRN<sub>L.h.</sub> verschob sich nach Absenken des pH-Wertes und der NaCl-Konzentration in Richtung der Dimere. Die Restriktion von D39R<sub>L.h.</sub> in Konformation 1 korrelierte mit der Stabilisierung als Monomer, wohingegen eine vergleichbare strukturelle Limitierung die

pH-induzierte Dimerisierung von A69R<sub>L.h.</sub> nicht beeinflusste. Unabhängig voneinander durchgeführte MALS-Messungen und NMR-Relaxationsexperimente zeigten eine nahezu pH-unabhängige Stabilisierung der Variante E114Q<sub>L.h.</sub> im dimeren Zustand. Dies bestätigt die für die homologe Variante von NRN<sub>E.a.</sub> beschriebene Verbesserung der Bindungsenergie des Dimers und Absenkung der Dissoziationsrate um das Fünffache auf  $k_{off} = 0,06 \pm 0,01 \text{ s}^{-1}$ .<sup>189, 197, 198</sup>



**Abbildung 15: Schema der pH-induzierten Dimerisierung und Konformationsänderung von NRN1<sub>L.h.</sub>** **A)** Die bei physiologischem pH-Wert deprotonierten D39, E76, E81 (rote Kugeln) und E114 (orange) stabilisieren NRN1<sub>L.h.</sub> in der monomeren Konformation I. Eine sinkende NaCl-Konzentration ermöglicht über elektrostatische Interaktionen zwischen dem sauren und basischen Pol eine antiparallele Vorausrichtung der beiden Untereinheiten. Ein sinkender pH-Wert löst über die Protonierung von E114 (orange → weiß) die Ausbildung stabiler Dimere aus. **B)** Der strukturelle Übergang von Konformation I zu II, bei dem sich das W9 Richtung Oberfläche verschiebt (gelb → grün), wird über die schrittweise Protonierung der zusammengelagerten Carboxylatgruppen von D39, E76 und E81 (rot → weiß) kontrolliert. Modifiziert nach Bauer *et al.*<sup>221</sup>

Die Änderung der Tertiär- und Quartärstruktur von NRN1<sub>L.h.</sub> werden somit unabhängig voneinander von verschiedenen Aminosäuren der Domäne kontrolliert. Die beiden Untereinheiten werden antiparallel ausgerichtet, sobald die NaCl Konzentration absinkt und elektrostatische Interaktionen zwischen dem sauren und basischen Cluster ausgebildet



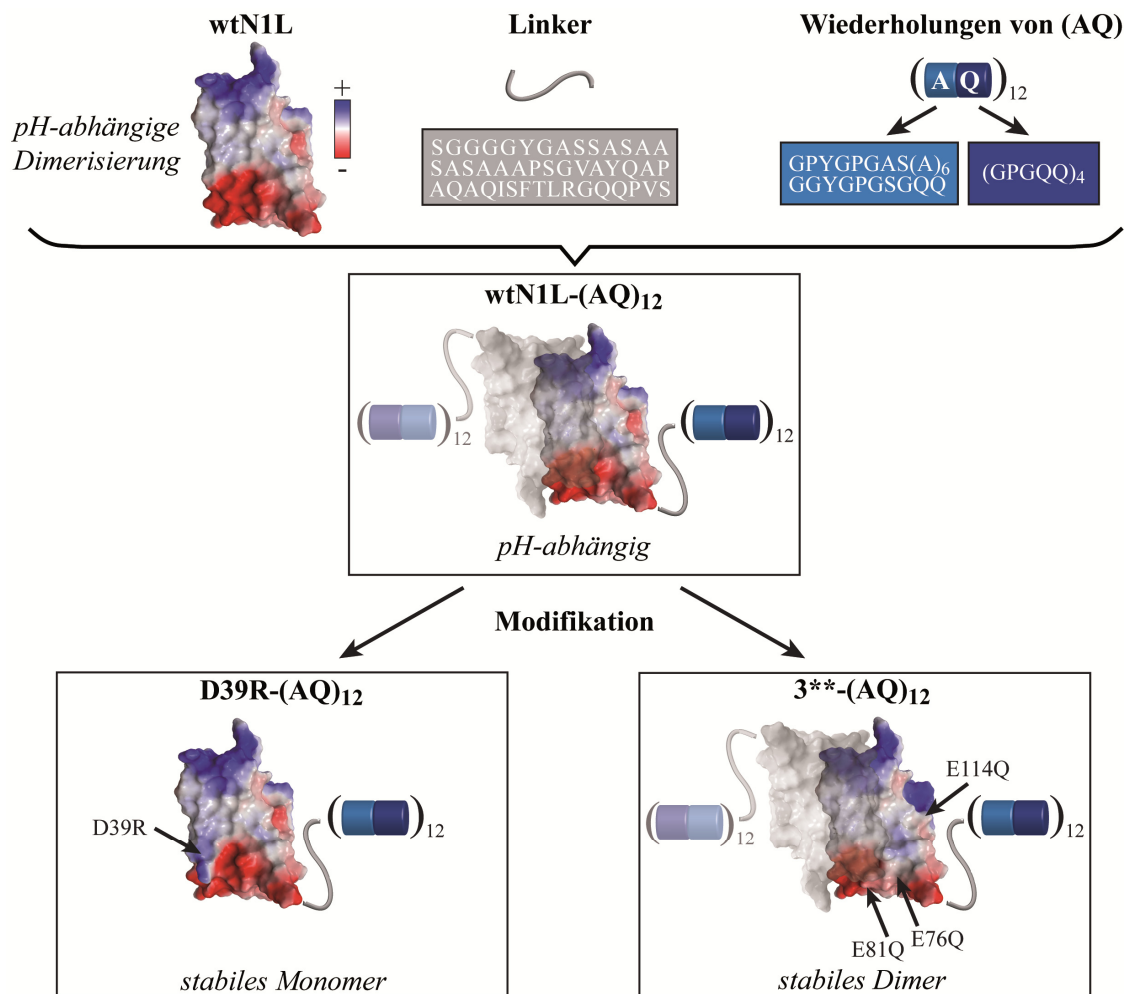
werden (Abbildung 15A). Die pH-induzierte Protonierung von E114 löst die Dimerbildung aus. Unabhängig von der Dimerisierung führt die Protonierung der zusammengelagerten Carboxylatgruppen von D39, E76 und E81 zu einer geringeren Ladungsabstoßung zwischen Helix 2 und 3. Als Folge dessen nimmt das Fünf-Helixbündel die engere Konformation II ein, in der das W9 Oberflächen-exponiert vorliegt (Abbildung 15B).

Der allgemeine Ablauf der Strukturänderung und Dimerisierung stimmt zwischen den NRN verschiedener MaSp überein und unterstreicht deren Bedeutung für die Assemblierung der Spidroine. Die exakten molekularen Mechanismen der orthologen Termini können jedoch auf Grund leicht unterschiedlicher Aminosäuresequenzen (Abbildung 6) geringfügig voneinander abweichen. Im Gegensatz zu NRN<sub>L.h.</sub> wird die Dimerisierung von NRN<sub>E.a.</sub> über die simultane Protonierung von E114 und E76 gesteuert, wohingegen E81 die finale Strukturänderung kontrolliert.<sup>191</sup> In dem untypisch asymmetrischen Dimerisierungsinterface von NRN<sub>N.c.</sub> aus *Nephila clavipes* MaSp1 sind die konservierten D39, K62 und E81 signifikant unterschiedlich angeordnet.<sup>200</sup> Als Folge dessen weist die Dimerisierung von NRN<sub>N.c.</sub> einen zusätzlichen pH-sensitiven Schritt auf, der auf der Wechselwirkung zwischen den weniger stark konservierten D16 und D52 beruht.<sup>200</sup> Bei der N-terminalen Domäne aus *Araneus ventricosus* MiSp (NRN<sub>A.v.</sub>) ist die Glutaminsäure an Position 81 durch ein Serin ersetzt (Abbildung 6), sodass die Funktion von dem nur in NRN<sub>A.v.</sub> vorkommenden E73 übernommen wird.<sup>180</sup> Das ansonsten konservierte E114, dessen Protonierung innerhalb weniger Millisekunden die Dimerbildung der MaSp und MiSp auslöst,<sup>186</sup> ist bei den NRN der Tubuliform-Spidroinen (bzw. Cylindriform) teilweise durch ein Arginin ersetzt. Im Gegensatz zu den MaSp, die sehr schnell für den regelmäßigen Einsatz als Netzrahmen und *Life-line* versponnen werden müssen, werden die für den Bau der Eierhülle benötigten Tubuliform-Spidroine lediglich wenige Male im Leben einer weiblichen Spinne verwendet. Die Charakterisierung der NRN der Tubuliform-Spidroinen sollte weitere Erkenntnisse über Ähnlichkeiten und Unterschiede zwischen den Dimerisierungsmechanismen verschiedener NRN liefern.

### 3.3 WECHSELWIRKUNGEN ZWISCHEN NRN UND KERNDOMÄNE

Zu Beginn dieser Arbeit stand bereits die Hypothese, dass NRN nicht für die Entstehung geordneter Seidenfäden benötigt wird, die Domäne jedoch den Übergang von löslichen zu assemblierten Spidroinen in einer pH- und Salz-abhängigen Weise steuert.<sup>219</sup> Erste Experimente mit rekombinanten Mini-Spidroinen (27 kDa) deuteten an, dass die Aggregation der Kerndomäne durch den monomeren Zustand von NRN<sub>E.a.</sub> unterdrückt und durch den dimeren beschleunigt wird.<sup>139</sup>





**Abbildung 16: Übersicht über die entworfenen Konstrukte aus NRN1<sub>L.h.</sub> und der rekombinanten Kerndomäne (AQ).** Das in Kapitel 3.2 beschriebene wtNRN1<sub>L.h.</sub> wurde über einen kurzen *Linker* (L) mit zwölf Wiederholungen der (AQ)-Module zu dem Konstrukt wtN1L-(AQ)<sub>12</sub> zusammengefügt. Die Variante D39R-(AQ)<sub>12</sub> wurde hinsichtlich der Stabilisierung im monomeren und die Variante 3\*\*-(AQ)<sub>12</sub>, in der drei saure Aminosäuren (E76, E81 und E114) neutralisiert wurden, hinsichtlich der Stabilisierung im dimeren Zustand analysiert. Die Struktur von wtNRN1<sub>L.h.</sub> wurde mit PyMOL dargestellt und modifiziert.<sup>141</sup> Modifiziert nach <sup>223</sup> mit Copyright © 2017 American Chemical Society.

Nachdem in Publikation I und II die wichtigsten Stellschrauben des Dimerisierungsprozesses von NRN1<sub>L.h.</sub> identifiziert wurden, konnte in Publikation III der Einfluss der Stabilisierung des monomeren bzw. dimeren Zustands von NRN1<sub>L.h.</sub> auf die Assemblierung einer rekombinanten Kerndomäne analysiert werden.<sup>223</sup> *Vice versa* wurde untersucht, wie die Dimerisierung von NRN1<sub>L.h.</sub> durch die Fusion mit einer im löslichen Zustand sperrigen Kerndomäne beeinträchtigt wird. Neben dem wtNRN1<sub>L.h.</sub>, dessen Monomer-Dimer Gleichgewicht sich entlang des Protonen- und Salz-Gradienten verschiebt, wurden erwartungsgemäß pH-unabhängige Varianten des Monomers (D39R) und Dimers (3\*\*) über einen kurzen *Linker* mit zwölf Wiederholungen des zuvor etablierten (AQ)-Moduls<sup>160</sup> zu N1L-(AQ)<sub>12</sub> fusioniert (Abbildung 16). Die Variante 3\*\* (E76Q-E81Q-

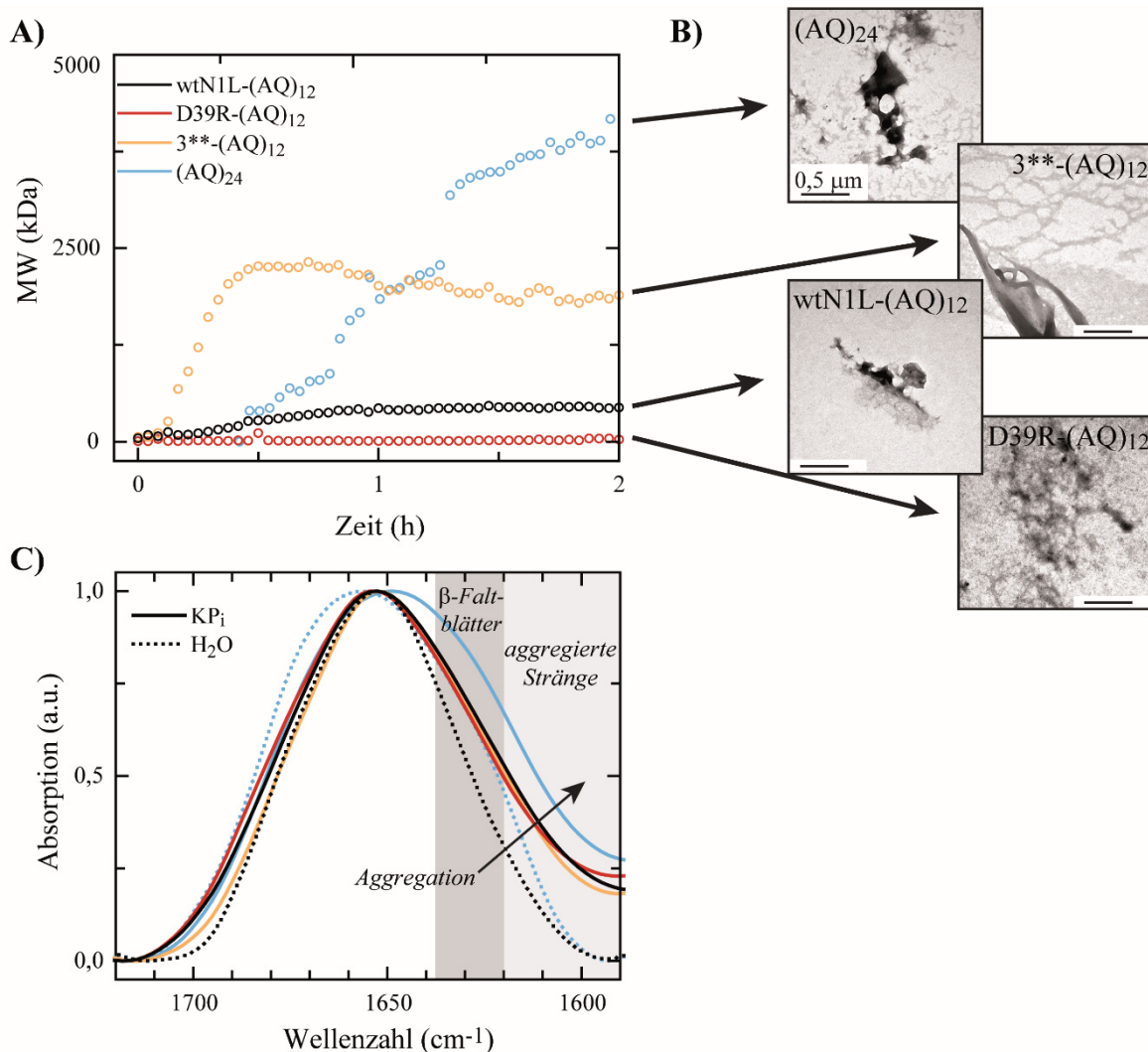
E114Q) wurde basierend auf Publikation I und zwischenzeitlich erschienen Veröffentlichungen ohne D39N entworfen, da die geladene Asparaginsäure 39 die antiparallele Dimerbindung über eine intermolekulare Salzbrücke mit Lysin 62 der zweiten Untereinheit unterstützt.<sup>139, 189, 191, 197, 198, 200-202, 221</sup>

Die als Fünf-Helix-Bündel vorliegende wtN1-Domäne und der unstrukturierte Abschnitt der (AQ)<sub>12</sub> Kerndomäne wurden durch Addition des *Linkers* und Fusion zu dem Konstrukt wtN1L-(AQ)<sub>12</sub> nicht beeinflusst. Ferner hatten der pH-Wert, die NaCl-Konzentration sowie die Umkehr (D39R) und Neutralisierung (3\*\*) der Ladungen konservierter D und E keinen Einfluss auf die Sekundärstruktur von NRN1<sub>L.h.</sub>, beeinflussten jedoch signifikant die Tertiär- und Quartärstruktur. Während wtN1L-(AQ)<sub>12</sub> pH-abhängig dimerisierte und zwischen Konformation I und II wechselte, wurden erwartungsgemäß 3\*\*-(AQ)<sub>12</sub> im dimeren Zustand stabilisiert und D39R-(AQ)<sub>12</sub> als nahezu pH-unabhängiges Monomer in Konformation I fixiert. Somit blieb der komplexe Dimerisierungsmechanismus von NRN trotz Fusion mit einer großen, unstrukturierten Kerndomäne erhalten.

Im nächsten Schritt wurde der Einfluss des monomeren und dimeren NRN auf die Kaliumphosphat-induzierte (KPi) Assemblierung der Spidroine mit Hilfe von MALD, Transmissionselektronenmikroskopie (TEM) und Fourier-Transform-Infrarotspektroskopie (FTIR) untersucht. Das NRN-freie (AQ)<sub>24</sub> aggregierte kontinuierlich in unstrukturierte Agglomerate mit hohem Molekulargewicht und  $\beta$ -Faltblattanteil (Abbildung 17A, B, C). Gemäß der suggerierten natürlichen Funktion wirkte die Addition der N-terminalen Domäne in wtN1L-(AQ)<sub>12</sub> der Aggregation der Kerndomäne entgegen (Abbildung 17A, B). Die Stabilisierung von D39R-(AQ)<sub>12</sub> im monomeren Zustand unterdrückte vollständig die Bildung von Aggregaten, wohingegen der stabile, dimere Zustand von 3\*\*-(AQ)<sub>12</sub> die Assemblierungskinetik beschleunigte und sowohl die Größe, als auch Ausrichtung der Agglomerate steigerte (Abbildung 17A, B). Die  $\alpha$ -helikalen Bereiche der NRN-Domäne blieben in den Aggregaten erhalten, hinderten jedoch die Kerndomäne nicht an der Ausbildung von  $\beta$ -Faltblattstrukturen (Abbildung 17C). Neben der KPi- wurde auch die Scherkraft-induzierte Bildung von fibrillären Assemblaten durch den monomeren Zustand von NRN unterdrückt und durch den dimeren Zustand forciert.

Zusammengefasst vermittelte der monomere Zustand von NRN1<sub>L.h.</sub> bei neutralem pH-Wert und in der Anwesenheit von NaCl die Löslichkeit der Spidroine, was deren Lagerung im Reservoir der Spinndrüse unterstützt und eine verfrühte Aggregation verhindert. Der intrinsische Protonengradient in der Spinndrüse bewirkt über die Protonierung von konservierten Glutamaten (E76, E81 und E114) die energetische Stabilisierung des dimeren

Zustands und unterstützt die  $\text{KPi}$ - und Scherkraft-induzierte Selbstassemblierung der Kerndomäne in Fibrillen und Fasern. Diese Studie bestätigte das zuvor beschriebene Verhalten von rekombinanten Mini-Spidroinen, bei denen das homologe  $\text{NRN}_{E.a.}$  die vorzeitige Aggregation bei neutralem pH-Wert unterdrückte und die gerichtete Selbstassemblierung bei saurem pH-Wert beschleunigte.<sup>139</sup> Die Wichtigkeit der NRN Domäne für die Spidroinassemblierung wird durch die Tatsache unterstrichen, dass bei der Herstellung artifizieller Spinnenseidenfäden die versponnenen Spidroine die  $\text{NRN}_{L.h.}$ -Domäne benötigen, um die hohe Zähigkeit des natürlichen *Dragline*-Seidenfadens zu erreichen.<sup>173, 229</sup>



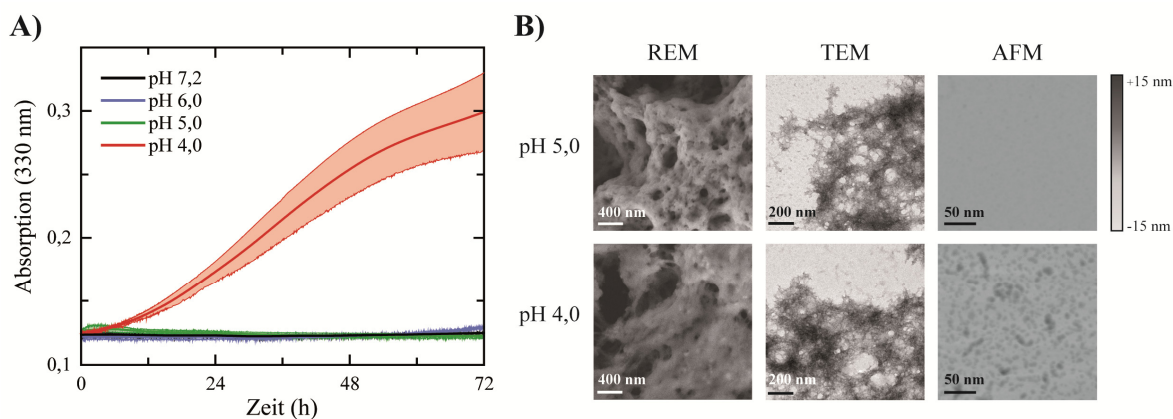
**Abbildung 17:  $\text{KPi}$ -induzierte Aggregation der Varianten von  $\text{N1L}-(\text{AQ})_{12}$ .** **A)** Die Aggregation wurde durch Zugabe von  $\text{KPi}$  ausgelöst und der zeitabhängige Verlauf der Molekulargewichtsänderung mit Hilfe von MALS bestimmt. **B)** Die gebildeten Agglomerate wurden mit TEM detektiert. **C)** Die FTIR Spektren verdeutlichen die  $\text{KPi}$ -induzierte Bildung von  $\beta$ -Faltblättern (dunkelgrau) und Aggregaten (hellgrau). Modifiziert nach Bauer *et al.*<sup>223</sup> mit Copyright © 2017 American Chemical Society.

### 3.4 STRUKTURELLE UND FUNKTIONELLE CHARAKTERISIERUNG DER NICHT-REPETITIVEN C-TERMINALEN DOMÄNE NRC

Der dreiteilige Aufbau der Spidroine wird mit der Anwesenheit der nicht-repetitiven C-terminalen (NRC) Domäne komplettiert, deren Fusion an die rekombinanten Spidroinen zudem die mechanischen Eigenschaften der artifiziellen Seidenfäden verbesserte.<sup>173</sup> Der Domäne wurde eine Funktion als mechano-sensitiver Schalter zugewiesen, der zunächst den löslichen Zustand der Spidroine stabilisiert und nach Einwirken von Scherkräften über die Exponierung hydrophober Bereiche die Assemblierung der repetitiven Module in  $\beta$ -Faltblatt-reiche Strukturen beschleunigt und koordiniert.<sup>127, 138</sup> Darüber hinaus wurde bei NRC<sub>A.d.</sub>,<sup>138</sup> NRC<sub>A.v.</sub> (NRC aus *A. ventricosus* MiSp),<sup>131</sup> und NRC<sub>N.c.</sub> (NRC aus *N. clavipes* MaSp1)<sup>209</sup> bei einem nicht-physiologischen pH-Wert von  $\leq$  pH 5,5 ein *molten globule* Zustand mit partiell exponierten, hydrophoben Bereichen detektiert, aus dem Amyloid-ähnliche Fasern mit erhöhtem  $\beta$ -Faltblattanteil hervorgingen (siehe 1.7.5). Um zu überprüfen, ob die Ausbildung amyloider Fibrillen eine Bedeutung für die natürliche Spidroinassemblierung hat, wurde in Publikation IV NRC1<sub>L.h.</sub> aus *Latrodectus hesperus* MaSp1 über einen breiten pH-Bereich charakterisiert.<sup>224</sup>

Während NRN1<sub>L.h.</sub> bei niedrigem pH-Wert die stabilisierte Konformation II einnahm (siehe 3.2), wurden die Sekundär- und Tertiärstruktur von NRC1<sub>L.h.</sub> bei pH  $\leq$  5,0 destabilisiert und hydrophobe Bereiche exponiert. Dies bestätigte die in der Literatur beschriebene, partielle Entfaltung der homologen NRC<sub>A.d.</sub>, NRC<sub>A.v.</sub>, und NRC<sub>N.c.</sub> bei pH  $<$  5,0,<sup>131, 138, 209</sup> sodass die Destabilisierung von NRC im extrem niedrigen pH-Bereich eine zwischen orthologen Spidroinen konservierte Funktion darzustellen scheint. Das Fünf-Helix-Bündel wird bei physiologischem pH-Wert über intramolekulare Salzbrücken zwischen hoch-konservierten sauren und basischen Aminosäuren stabilisiert (NRC<sub>A.d.</sub>: R43–D93 und R52–E101, siehe Abbildung 10)<sup>138</sup>. Bei einem nicht-physiologischen pH-Wert von 4,0 werden vermutlich die Carboxylatgruppen der sauren Aminosäureseitenketten (theoretischer pK<sub>S</sub> von D = 4,0; E = 4,4; siehe Tabelle 2) protoniert, sodass die Salzbrücken aufbrechen, die Sekundär- und Tertiärstruktur destabilisiert und hydrophobe Bereiche exponiert werden. Eine über mehrere Tage ansteigende Trübung signalisierte zudem die langsame Aggregation des bei pH 4,0 destabilisierten NRC1<sub>L.h.</sub> (Abbildung 18A). Die Aggregation von NRC1<sub>L.h.</sub> wies eine deutlich langsamere Kinetik auf als die innerhalb von Millisekunden ablaufende, natürliche Spidroinassemblierung.<sup>56</sup> Über die Bindung von Thioflavin T und Kongorot wurden in den NRC1<sub>L.h.</sub>-Aggregaten  $\beta$ -Faltblattstrukturen nachgewiesen, die jedoch nicht den  $\beta$ -Faltblattanteil, die Rigidität und den Grad der Ausrichtung von eADF4-Spidroinfibrillen<sup>160</sup> erreichten. Anstatt definierte Fibrillen auszubilden, aggregierte NRC1<sub>L.h.</sub> in sphärische und

10 nm große Agglomerate, die sich zu einer porösen und heterogenen Matrix mit einer Größe von mehreren hundert nm zusammenlagerten (Abbildung 18B). Die Aggregate von NRC1<sub>L.h.</sub> unterschieden sich signifikant von den gestreckten und Amyloid-ähnlichen Fibrillen der homologen NRC<sub>A.v.</sub> und NRC<sub>N.c.</sub>,<sup>131, 209</sup> die somit eine eher Spidroin-spezifische Besonderheit als ein NRC-konserviertes Verhalten darstellen.

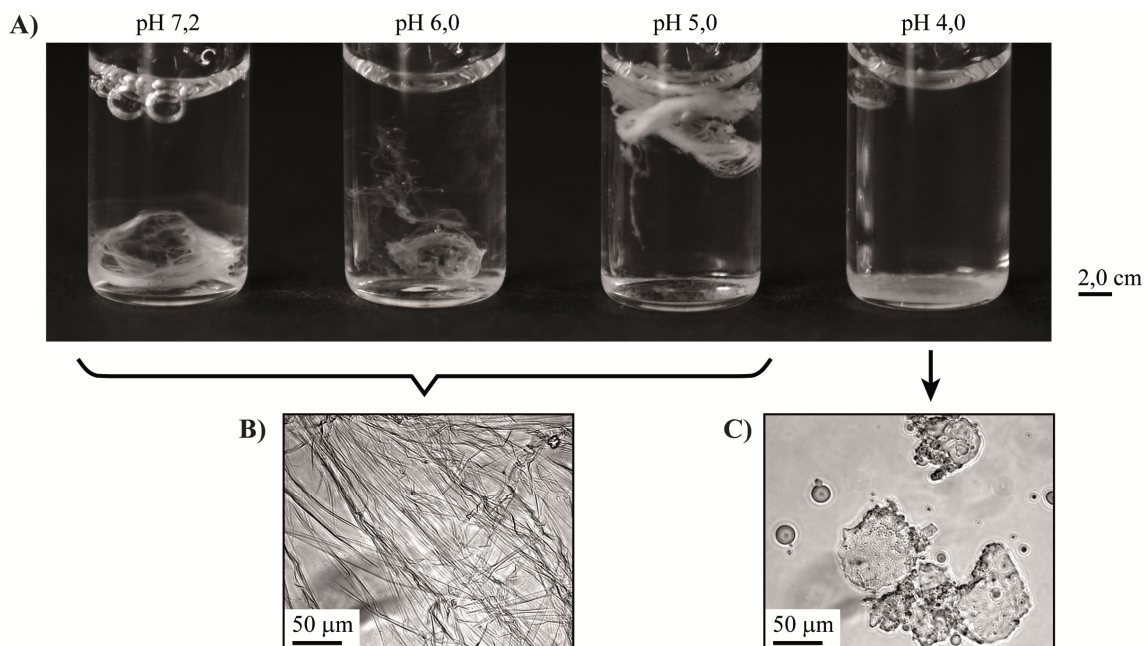


**Abbildung 18: pH-abhängige Strukturanalyse von NRC1<sub>L.h.</sub>** A) Die zeitabhängige Absorption (Trübung) deutete auf die Aggregation von NRC1<sub>L.h.</sub> bei pH 4,0 hin. B) Die Morphologie der bei pH 5,0 und pH 4,0 entstanden NRC1<sub>L.h.</sub>-Aggregate wurde mittels Rasterelektronenmikroskop (REM), TEM und Rasterkraftmikroskopie (AFM) detektiert. Modifiziert nach <sup>224</sup> mit Copyright © 2017 American Chemical Society.

### 3.5 WECHSELWIRKUNGEN ZWISCHEN NRC UND KERNDOMÄNE

Der Einfluss von NRC, das auf Grund der Destabilisierung bei niedrigem pH-Wert zum Aggregieren tendiert, auf die Selbstassemblierung der repetitiven Module wurde in Publikation IV analysiert. Dafür wurde NRC1<sub>L.h.</sub> mit der (AQ)<sub>12</sub> Kerndomäne zu dem Konstrukt (AQ)<sub>12</sub>-NRC1 fusioniert.<sup>224</sup> Die natürliche Funktion von NRC als Scherkraftsensitiver Schalter, der die Selbstassemblierung der Spidroine initiiert und gleichzeitig die repetitiven Module ausrichtet,<sup>138</sup> wurde anschließend pH-abhängig untersucht. Erwartungsgemäß unterstützte NRC1<sub>L.h.</sub> im physiologischen pH-Bereich zwischen pH 7,2 und pH 5,0 die Assemblierung von (AQ)<sub>12</sub>-NRC1 in Millimeter große, stark geordnete Fibrillenbündel (Abbildung 19A, B). Dahingegen ging bei dem nicht-physiologischen pH 4,0 die natürliche Funktion von NRC1<sub>L.h.</sub> verloren, sodass (AQ)<sub>12</sub>-NRC1 unter Scherbelastung unstrukturierte und nicht-faserartige Aggregate bildete (Abbildung 19C). Die Säure-induzierte, unspezifische Aggregation wurde gleichermaßen für ein Konstrukt aus (AQ)<sub>12</sub> und dem homologen NRC<sub>A.d.</sub> (BLAST: 81 % Ähnlichkeit zu NRC1<sub>L.h.</sub>)<sup>230</sup> detektiert, als Folge dessen der Funktionsverlust von NRC bei niedrigem pH ein konserviertes aber unnatürliches Verhalten darzustellen scheint. Die bei nicht-physiologischem pH-Wert ablaufende Aggregation von NRC in  $\beta$ -Faltblatt-reiche Strukturen dient nicht als Keim, der

die Selbstassemblierung der Spidroine beschleunigt, sondern unterscheidet sich klar von der dynamischen, kontrollierten und reversiblen Spidroinassemblierung in geordnete Fibrillenbündel.



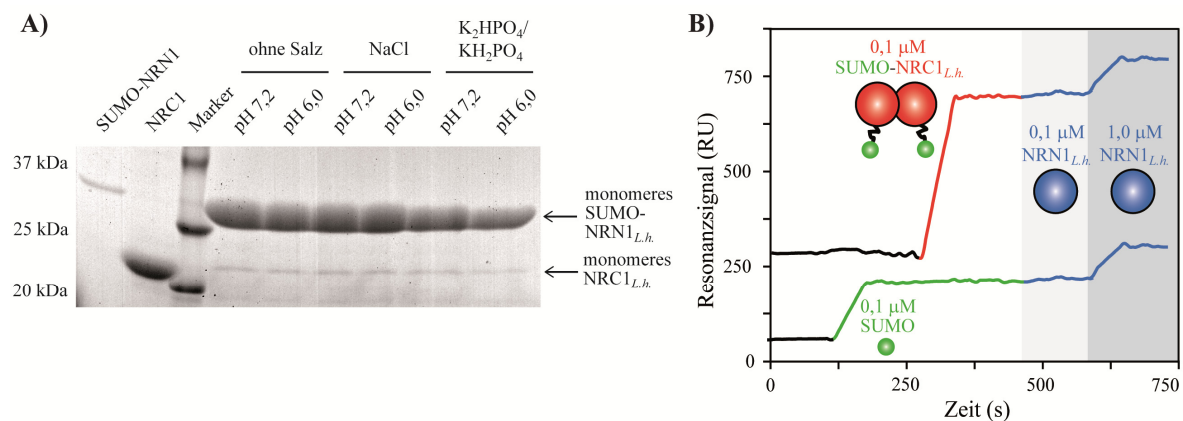
**Abbildung 19: Die Scherkraft-induzierte Assemblierung von (AQ)<sub>12</sub>-NRC1 wurde pH-abhängig untersucht.** A) Die Scherkraft-induzierten Aggregate von (AQ)<sub>12</sub>-NRC1 offenbarten zwischen pH 7,2 und pH 5,0 fibrilläre Strukturen, die bei pH 4,0 nicht gebildet wurden. B) Aufnahmen der Durchlichtmikroskopie bestätigten die Bildung von (AQ)<sub>12</sub>-NRC1 Fibrillenbündel in zwischen pH 7,2 - 5,0. C) Bei pH 4,0 bildete (AQ)<sub>12</sub>-NRC1 unstrukturierte und kleinere Aggregate. Modifiziert nach Bauer *et al.*<sup>224</sup> mit Copyright © 2017 American Chemical Society.

### 3.6 WECHSELWIRKUNGEN ZWISCHEN TERMINALEN DOMÄNEN

Die terminalen Domänen NRN und NRC zeigen einige Gemeinsamkeiten wie beispielsweise ihre hohe Konservierung zwischen verschiedenen Spidroinen (Abbildung 6, Abbildung 9), die Faltung in eine Fünf-Helix-Bündel Struktur (Abbildung 7, Abbildung 10), die Dimerisierung (Abbildung 10), die Prävention einer verfrühten Koaleszenz während der Lagerung der Spidroine (siehe 1.6 und 1.7.5), sowie die Initiierung der Assemblierung der Kerndomäne (siehe 3.3 und 3.5). Um zu überprüfen, ob sich die Strukturen der helikalen Termini gegenseitig beeinflussen, wurde in Publikation IV die Struktur von NRN1<sub>L.h.</sub> und NRC1<sub>L.h.</sub> in heterogenen Proteinmischungen analysiert.<sup>224</sup> Die  $\alpha$ -helikale Sekundärstruktur von NRN1<sub>L.h.</sub> und NRC1<sub>L.h.</sub> blieb in dem heterogenen Proteingemisch bei physiologischem pH erhalten. Die bei pH 4,0 auftretende Exponierung hydrophober NRC1<sub>L.h.</sub>-Bereiche und Bildung von  $\beta$ -Faltblattstrukturen (siehe 3.4) löste nicht die Aggregation von NRN1<sub>L.h.</sub> aus. Allerdings nahm die thermische Stabilität der Sekundärstruktur von NRN1<sub>L.h.</sub> und NRC1<sub>L.h.</sub> im Proteingemisch bei pH 6,0 ab. Um aufzudecken, ob die molekularen Schalter während



der Assemblierung individuell oder im Zusammenspiel agieren, wurden diese im nächsten Schritt bezüglich möglicher Wechselwirkungen analysiert. In einem *pull-down assay* trat bei Lösungsmittelbedingungen, die in der Spinndrüse die Lagerung bzw. Assemblierung der Spidroine unterstützen, keine stabile Interaktion zwischen  $\text{NRC1}_{L.h.}$  und dem immobilisierten  $\text{SUMO-NRN1}_{L.h.}$  auf (Abbildung 20A). Zudem wurde mittels hochsensitiven *surface plasmon resonance* (englisch, SPR, Oberflächenplasmonenresonanz) Spektroskopiemessungen keine kurzzeitigen und dynamischen Wechselwirkungen zwischen den Termini detektiert (Abbildung 20B). Das Fehlen von intermolekularen Interaktionen zwischen den terminalen Domänen lässt einen doppelten Sicherheitsmechanismus entstehen, in dem die NRN-Domäne Ionen-sensitiv und unabhängig davon die NRC-Domäne nach Einwirken von Scherkräften den Übergang vom löslichen zum assemblierten Zustand kontrollieren.



**Abbildung 20: Analyse von heterogenen Proteingemischen aus  $\text{NRN1}_{L.h.}$  und  $\text{NRC1}_{L.h.}$ .** A) Die SDS-PAGE Analyse der im *pull-down assay* gebundenen Proteine signalisierte das Ausbleiben stabiler Interaktionen zwischen  $\text{SUMO-NRN1}_{L.h.}$  und  $\text{NRC1}_{L.h.}$ . B) Die SPR spektroskopische Analyse offenbarte eine extrem niedrige Affinität des verschieden konzentrierten  $\text{NRN1}_{L.h.}$  zu  $\text{SUMO-NRC1}_{L.h.}$  und zu  $\text{SUMO-tag}$ . Modifiziert nach *Bauer et al.*<sup>224</sup> mit Copyright © 2017 American Chemical Society.

#### 4. LITERATURVERZEICHNIS

- 1 Whitesides, G. M. & Grzybowski, B. Self-assembly at all scales. *Science* **2002**, 295, (5564), 2418-2421.
- 2 Chan, H. S. & Dill, K. A. The Protein Folding Problem. *Physics Today* **1993**, 46, (2), 24-32.
- 3 Kaplan, D. L. Fibrous proteins - silk as a model system. *Polymer Degradation and Stability* **1998**, 59, (1-3), 25-32.
- 4 Dobson, C. M. Protein folding and misfolding. *Nature* **2003**, 426, (6968), 884-890.
- 5 Andreeva, A., Howorth, D., Chothia, C., Kulesha, E. & Murzin, A. G. SCOP2 prototype: a new approach to protein structure mining. *Nucleic Acids Research* **2014**, 42, (D1), D310-D314.
- 6 Dyson, H. J. & Wright, P. E. Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology* **2005**, 6, (3), 197-208.
- 7 Saitô, H., Ando, I. & Naito, A. Fibrous Proteins. In *Solid state NMR spectroscopy for biopolymers principles and applications*, Springer: Dordrecht, **2006**, 241-287.
- 8 Foo, C. W. P. & Kaplan, D. L. Genetic engineering of fibrous proteins: spider dragline silk and collagen. *Advanced Drug Delivery Reviews* **2002**, 54, (8), 1131-1143.
- 9 Scheibel, T. Protein fibers as performance proteins: new technologies and applications. *Current Opinion in Biotechnology* **2005**, 16, (4), 427-433.
- 10 Barbu, E. & Joly, M. The Globular-Fibrous Protein Transformation. *Discussions of the Faraday Society* **1953**, (13), 77-93.
- 11 Shoulders, M. D. & Raines, R. T. Collagen Structure and Stability. *Annual Review of Biochemistry* **2009**, 78, 929-958.
- 12 Pollard, T. D., Blanchoin, L. & Mullins, R. D. Molecular mechanisms controlling actin filament dynamics in nonmuscle cells. *Annual Review of Biophysics and Biomolecular Structure* **2000**, 29, 545-576.
- 13 Condeelis, J. How is actin polymerization nucleated in vivo? *Trends in Cell Biology* **2001**, 11, (7), 288-293.
- 14 Welch, M. D. & Mullins, R. D. Cellular control of actin nucleation. *Annual Review of Cell and Developmental Biology* **2002**, 18, 247-288.
- 15 Cooper, J. A. & Schafer, D. A. Control of actin assembly and disassembly at filament ends. *Current Opinion in Cell Biology* **2000**, 12, (1), 97-103.
- 16 Kielty, C. M., Sherratt, M. J. & Shuttleworth, C. A. Elastic fibres. *Journal of Cell Science* **2002**, 115, (14), 2817-2828.
- 17 Muiznieks, L. D. & Keeley, F. W. Molecular assembly and mechanical properties of the extracellular matrix: A fibrous protein perspective. *Biochimica Et Biophysica Acta-Molecular Basis of Disease* **2013**, 1832, (7), 866-875.
- 18 Bellingham, C. M., Woodhouse, K. A., Robson, P., Rothstein, S. J. & Keeley, F. W. Self-aggregation characteristics of recombinantly expressed human elastin polypeptides. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* **2001**, 1550, (1), 6-19.
- 19 Toonkool, P., Jensen, S. A., Maxwell, A. L. & Weiss, A. S. Hydrophobic domains of human tropoelastin interact in a context-dependent manner. *Journal of Biological Chemistry* **2001**, 276, (48), 44575-44580.
- 20 Herrmann, H., Wedig, T., Porter, R. M., Lane, E. B. & Aebi, U. Characterization of early assembly intermediates of recombinant human keratins. *Journal of Structural Biology* **2002**, 137, (1-2), 82-96.
- 21 Lee, C. H., Kim, M. S., Chung, B. M., Leahy, D. J. & Coulombe, P. A. Structural basis for heteromeric assembly and perinuclear organization of keratin filaments. *Nature Structural & Molecular Biology* **2012**, 19, (7), 707-+.



- 22 Lee, C. H. & Coulombe, P. A. Self-organization of keratin intermediate filaments into cross-linked networks. *Journal of Cell Biology* **2009**, 186, (3), 409-421.
- 23 Silver, F. H., Freeman, J. W. & Seehra, G. P. Collagen self-assembly and the development of tendon mechanical properties. *Journal of Biomechanics* **2003**, 36, (10), 1529-1553.
- 24 Kadler, K. E., Holmes, D. F., Trotter, J. A. & Chapman, J. A. Collagen fibril formation. *Biochemical Journal* **1996**, 316, 1-11.
- 25 Fitch, J. M., Gross, J., Mayne, R., Johnsonwint, B. & Linsenmayer, T. F. Organization of Collagen Type-I and Type-V in the Embryonic Chicken Cornea - Monoclonal-Antibody Studies. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **1984**, 81, (9), 2791-2795.
- 26 O'Brien, E. J. & Bennett, P. M. Structure of Straight Flagella from a Mutant Salmonella. *Journal of Molecular Biology* **1972**, 70, (1), 133-&.
- 27 Samatey, F. A., Imada, K., Nagashima, S., Vonderviszt, F., Kumasaka, T., Yamamoto, M. & Namba, K. Structure of the bacterial flagellar protofilament and implications for a switch for supercoiling. *Nature* **2001**, 410, (6826), 331-337.
- 28 Rambaran, R. N. & Serpell, L. C. Amyloid fibrils Abnormal protein assembly. *Prion* **2008**, 2, (3), 112-117.
- 29 Fitzpatrick, A. W. P., Debelouchina, G. T., Bayro, M. J., Clare, D. K., Caporini, M. A., Bajaj, V. S., Jaroniec, C. P., Wang, L. C., Ladizhansky, V., Muller, S. A., MacPhee, C. E., Waudby, C. A., Mott, H. R., De Simone, A., Knowles, T. P. J., Saibil, H. R., Vendruscolo, M., Orlova, E. V., Griffin, R. G. & Dobson, C. M. Atomic structure and hierarchical assembly of a cross-beta amyloid fibril. *Proceedings of the National Academy of Sciences of the United States of America* **2013**, 110, (14), 5468-5473.
- 30 Straub, J. E. & Thirumalai, D. Toward a molecular theory of early and late events in monomer to amyloid fibril formation. *Annual Review of Physical Chemistry* **2011**, 62, 437-463.
- 31 Kelly, J. W. The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Current Opinion in Structural Biology* **1998**, 8, (1), 101-106.
- 32 Lomakin, A., Chung, D. S., Benedek, G. B., Kirschner, D. A. & Teplow, D. B. On the nucleation and growth of amyloid beta-protein fibrils: Detection of nuclei and quantitation of rate constants. *Proceedings of the National Academy of Sciences of the United States of America* **1996**, 93, (3), 1125-1129.
- 33 Esler, W. P., Stimson, E. R., Ghilardi, J. R., Vinters, H. V., Lee, J. P., Mantyh, P. W. & Maggio, J. E. In vitro growth of Alzheimer's disease beta-amyloid plaques displays first-order kinetics. *Biochemistry* **1996**, 35, (3), 749-757.
- 34 Harper, J. D., Wong, S. S., Lieber, C. M. & Lansbury, P. T. Observation of metastable A beta amyloid protofibrils by atomic force microscopy. *Chemistry & Biology* **1997**, 4, (2), 119-125.
- 35 Walsh, D. M., Lomakin, A., Benedek, G. B., Condron, M. M. & Teplow, D. B. Amyloid beta-protein fibrillogenesis - Detection of a protofibrillar intermediate. *Journal of Biological Chemistry* **1997**, 272, (35), 22364-22372.
- 36 Craig, C. L. Evolution of arthropod silks. *Annual Review of Entomology* **1997**, 42, 231-267.
- 37 Knight, D. P. & Vollrath, F. Biological liquid crystal elastomers. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **2002**, 357, (1418), 155-163.
- 38 Rudall, K. M. & Kenchington, W. Arthropod Silks - Problem of Fibrous Proteins in Animal Tissues. *Annual Review of Entomology* **1971**, 16, 73-96.

- 39 Guerette, P. A., Ginzinger, D. G., Weber, B. H. F. & Gosline, J. M. Silk properties determined by gland-specific expression of a spider fibroin gene family. *Science* **1996**, 272, (5258), 112-115.
- 40 Simmons, A. H., Michal, C. A. & Jelinski, L. W. Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk. *Science* **1996**, 271, (5245), 84-87.
- 41 Xu, M. & Lewis, R. V. Structure of a protein superfiber – Spider dragline silk. *Proceedings of the National Academy of Sciences of the United States of America* **1990**, 87, (18), 7120-7124.
- 42 Sutherland, T. D., Young, J. H., Weisman, S., Hayashi, C. Y. & Merritt, D. J. Insect silk: One name, many materials. *Annual Review of Entomology* **2010**, 55, 171-188.
- 43 Tanaka, K., Mori, K. & Mizuno, S. Immunological Identification of the Major Disulfide-Linked Light Component of Silk Fibroin. *Journal of Biochemistry* **1993**, 114, (1), 1-4.
- 44 Miller, L. D., Putthanarat, S., Eby, R. K. & Adams, W. W. Investigation of the nanofibrillar morphology in silk fibers by small angle X-ray scattering and atomic force microscopy. *International Journal of Biological Macromolecules* **1999**, 24, (2-3), 159-165.
- 45 He, Y. X., Zhang, N. N., Li, W. F., Jia, N., Chen, B. Y., Zhou, K., Zhang, J. H., Chen, Y. X. & Zhou, C. Z. N-Terminal Domain of Bombyx mori Fibroin Mediates the Assembly of Silk in Response to pH Decrease. *Journal of Molecular Biology* **2012**, 418, (3-4), 197-207.
- 46 Sehna, F. & Zurovec, M. Construction of silk fiber core in Lepidoptera. *Biomacromolecules* **2004**, 5, (3), 666-674.
- 47 Vollrath, F. Strength and structure of spiders' silks. *Journal of Biotechnology* **2000**, 74, (2), 67-83.
- 48 Stauffer, S. L., Coguille, S. L. & Lewis, R. V. Comparison of physical properties of 3 silks from *Nephila clavipes* and *Araneus gemmoides*. *Journal of Arachnology* **1994**, 22, (1), 5-11.
- 49 Vollrath, F. & Knight, D. P. Liquid crystalline spinning of spider silk. *Nature* **2001**, 410, (6828), 541-548.
- 50 Vollrath, F. & Porter, D. Spider silk as archetypal protein elastomer. *Soft Matter* **2006**, 2, (5), 377-385.
- 51 Gosline, J. M., Demont, M. E. & Denny, M. W. The Structure and Properties of Spider Silk. *Endeavour* **1986**, 10, (1), 37-43.
- 52 Shear, W. A., Palmer, J. M., Coddington, J. A. & Bonamo, P. M. A Devonian Spinneret: Early Evidence of Spiders and Silk Use. *Science* **1989**, 246, (4929), 479-481.
- 53 Shultz, J. W. The Origin of the Spinning Apparatus in Spiders. *Biological Reviews* **1987**, 62, (2), 89-113.
- 54 Gosline, J. M., Guerette, P. A., Ortlepp, C. S. & Savage, K. N. The mechanical design of spider silks: From fibroin sequence to mechanical function. *Journal of Experimental Biology* **1999**, 202, (23), 3295-3303.
- 55 Kohler, T. & Vollrath, F. Thread Biomechanics in the 2 Orb-Weaving Spiders *Araneus-Diadematus* (Araneae, Araneidae) and *Uloborus-Walckenaerius* (Araneae, Uloboridae). *Journal of Experimental Zoology* **1995**, 271, (1), 1-17.
- 56 Vollrath, F., Madsen, B. & Shao, Z. Z. The effect of spinning conditions on the mechanics of a spider's dragline silk. *Proceedings of the Royal Society of London Series B-Biological Sciences* **2001**, 268, (1483), 2339-2346.
- 57 Hayashi, C. Y. & Lewis, R. V. Molecular architecture and evolution of a modular spider silk protein gene. *Science* **2000**, 287, (5457), 1477-1479.

- 58 Garb, J. E. & Hayashi, C. Y. Modular evolution of egg case silk genes across orb-weaving spider superfamilies. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, 102, (32), 11379-11384.
- 59 Gatesy, J., Hayashi, C., Motriuk, D., Woods, J. & Lewis, R. Extreme diversity, conservation, and convergence of spider silk fibroin sequences. *Science* **2001**, 291, (5513), 2603-2605.
- 60 Tian, M. Z. & Lewis, R. V. Molecular characterization and evolutionary study of spider tubuliform (eggcase) silk protein. *Biochemistry* **2005**, 44, (22), 8006-8012.
- 61 Brooks, A. E., Steinkraus, H. B., Nelson, S. R. & Lewis, R. V. An investigation of the divergence of major ampullate silk fibers from *Nephila clavipes* and *Argiope aurantia*. *Biomacromolecules* **2005**, 6, (6), 3095-3099.
- 62 Babb, P. L., Lahens, N. F., Correa-Garhwal, S. M., Nicholson, D. N., Kim, E. J., Hogenesch, J. B., Kuntner, M., Higgins, L., Hayashi, C. Y., Agnarsson, I. & Voight, B. F. The *Nephila clavipes* genome highlights the diversity of spider silk genes and their complex expression. *Nature genetics* **2017**.
- 63 Eisoldt, L., Smith, A. & Scheibel, T. Decoding the secrets of spider silk. *Materials Today* **2011**, 14, (3), 80-86.
- 64 Gosline, J., Lillie, M., Carrington, E., Guerette, P., Ortlepp, C. & Savage, K. Elastic proteins: biological roles and mechanical properties. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **2002**, 357, (1418), 121-132.
- 65 Gosline, J. M., Denny, M. W. & Demont, M. E. Spider Silk as Rubber. *Nature* **1984**, 309, (5968), 551-552.
- 66 Gosline, J. M., Pollak, C. C., Guerette, P. A., Cheng, A., Demont, M. E. & Denny, M. W. Elastomeric network models for the frame and viscid silks from the orb web of the spider *Araneus diadematus*. *Silk Polymers* **1994**, 544, 328-341.
- 67 Viney, C. Natural silks: archetypal supramolecular assembly of polymer fibres. *Supramolecular Science* **1997**, 4, (1-2), 75-81.
- 68 Emile, O., Le Floch, A. & Vollrath, F. Biopolymers: Shape memory in spider draglines. *Nature* **2006**, 440, (7084), 621-621.
- 69 Emile, O., Le Floch, A. & Vollrath, F. Time-resolved torsional relaxation of spider draglines by an optical technique. *Physical Review Letters* **2007**, 98, (16), 167402.
- 70 Knight, D. P. & Vollrath, F. Liquid crystals and flow elongation in a spider's silk production line. *Proceedings: Biological Sciences* **1999**, 266, (1418), 519-523.
- 71 Swanson, B. O., Blackledge, T. A., Summers, A. P. & Hayashi, C. Y. Spider dragline silk: Correlated and mosaic evolution in high-performance biological materials. *Evolution* **2006**, 60, (12), 2539-2551.
- 72 Swanson, B. O., Blackledge, T. A., Beltran, J. & Hayashi, C. Y. Variation in the material properties of spider dragline silk across species. *Applied Physics a-Materials Science & Processing* **2006**, 82, (2), 213-218.
- 73 Agnarsson, I., Kuntner, M. & Blackledge, T. A. Bioprospecting finds the toughest biological material: extraordinary silk from a giant riverine orb spider. *PLoS ONE* **2010**, 5, (9), e11234.
- 74 Blackledge, T. A., Summers, A. P. & Hayashi, C. Y. Gumfooted lines in black widow cobwebs and the mechanical properties of spider capture silk. *Zoology* **2005**, 108, (1), 41-46.
- 75 Hayashi, C. Y. & Lewis, R. V. Evidence from flagelliform silk cDNA for the structural basis of elasticity and modular nature of spider silks. *Journal of Molecular Biology* **1998**, 275, (5), 773-784.
- 76 Vollrath, F. & Edmonds, D. T. Modulation of the mechanical properties of spider silk by coating with water. *Nature* **1989**, 340, (6231), 305-307.

- 77 Vollrath, F. & Tillinghast, E. K. Glycoprotein glue beneath a spider webs aqueous coat. *Naturwissenschaften* **1991**, 78, (12), 557-559.
- 78 Vollrath, F. Spider silk: Thousands of nano-filaments and dollops of sticky glue. *Current Biology* **2006**, 16, (21), R925-R927.
- 79 Hu, X. Y., Yuan, J., Wang, X. D., Vasanthavada, K., Falick, A. M., Jones, P. R., La Mattina, C. & Vierra, C. A. Analysis of aqueous glue coating proteins on the silk fibers of the cob weaver, *Latrodectus hesperus*. *Biochemistry* **2007**, 46, (11), 3294-3303.
- 80 Hawthorn, A. C. & Opell, B. D. Evolution of adhesive mechanisms in cribellar spider prey capture thread: evidence for van der Waals and hygroscopic forces. *Biological Journal of the Linnean Society* **2002**, 77, (1), 1-8.
- 81 Hawthorn, A. C. & Opell, B. D. Van-der-Waals and hygroscopic forces of adhesion generated by spider capture threads. *Journal of Experimental Biology* **2003**, 206, (22), 3905-3911.
- 82 Colgin, M. A. & Lewis, R. V. Spider minor ampullate silk proteins contain new repetitive sequences and highly conserved non-silk-like "spacer regions". *Protein Science* **1998**, 7, (3), 667-672.
- 83 Dicko, C., Knight, D., Kenney, J. M. & Vollrath, F. Secondary structures and conformational changes in flagelliform, cylindrical, major, and minor ampullate silk proteins. Temperature and concentration effects. *Biomacromolecules* **2004**, 5, (6), 2105-2115.
- 84 Riekel, C. & Vollrath, F. Spider silk fibre extrusion: combined wide- and small-angle X-ray microdiffraction experiments. *International Journal of Biological Macromolecules* **2001**, 29, (3), 203-210.
- 85 Guinea, G. V., Elices, M., Plaza, G. R., Perea, G. B., Daza, R., Riekel, C., Agullo-Rueda, F., Hayashi, C., Zhao, Y. & Perez-Rigueiro, J. Minor ampullate silks from *Nephila* and *Argiope* spiders: Tensile properties and microstructural characterization. *Biomacromolecules* **2012**, 13, (7), 2087-2098.
- 86 Blackledge, T. A., Boutry, C., Wong, S. C., Baji, A., Dhinojwala, A., Sahni, V. & Agnarsson, I. How super is supercontraction? Persistent versus cyclic responses to humidity in spider dragline silk. *Journal of Experimental Biology* **2009**, 212, (13), 1980-1988.
- 87 Kovoov, J. & Zylberberg, L. Fine-Structural Aspects of Silk Secretion in a Spider (*Araneus-Diadematus*) .1. Elaboration in the Pyriform Glands. *Tissue & Cell* **1980**, 12, (3), 547-556.
- 88 Geurts, P., Zhao, L., Hsia, Y., Gnesa, E., Tang, S., Jeffery, F., La Mattina, C., Franz, A., Larkin, L. & Vierra, C. Synthetic spider silk fibers spun from pyriform spidroin 2, a glue silk protein discovered in orb-weaving spider attachment discs. *Biomacromolecules* **2010**, 11, (12), 3495-3503.
- 89 Blasingame, E., Tuton-Blasingame, T., Larkin, L., Falick, A. M., Zhao, L., Fong, J., Vaidyanathan, V., Visperas, A., Geurts, P., Hu, X. Y., La Mattina, C. & Vierra, C. Pyriform Spidroin 1, a Novel Member of the Silk Gene Family That Anchors Dragline Silk Fibers in Attachment Discs of the Black Widow Spider, *Latrodectus hesperus*. *Journal of Biological Chemistry* **2009**, 284, (42), 29097-29108.
- 90 Hajer, J. & Rehakova, D. Spinning activity of the spider *Trogloneta granulum* (Araneae, Mysmenidae): web, cocoon, cocoon handling behaviour, draglines and attachment discs. *Zoology* **2003**, 106, (3), 223-231.
- 91 Bittencourt, D., Souto, B. M., Verza, N. C., Vineck, F., Dittmar, K., Silva, P. I., Andrade, A. C., da Silva, F. R., Lewis, R. V. & Rech, E. L. Spidroins from the Brazilian spider *Nephilengys cruentata* (Araneae: Nephilidae). *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology* **2007**, 147, (4), 597-606.

- 92 Hu, X. Y., Lawrence, B., Kohler, K., Falick, A. M., Moore, A. M. F., McMullen, E., Jones, P. R. & Vierra, C. Araneoid egg case silk: A fibroin with novel ensemble repeat units from the black widow spider, *Latrodectus hesperus*. *Biochemistry* **2005**, 44, (30), 10020-10027.
- 93 Huang, W., Lin, Z., Sin, Y. M., Li, D., Gong, Z. & Yang, D. Characterization and expression of a cDNA encoding a tubuliform silk protein of the golden web spider *Nephila antipodiana*. *Biochimie* **2006**, 88, (7), 849-858.
- 94 Hu, X. Y., Kohler, K., Falick, A. M., Moore, A. M. F., Jones, P. R., Sparkman, O. D. & Vierra, C. Egg case protein-1 - A new class of silk proteins with fibroin-like properties from the spider *Latrodectus hesperus*. *Journal of Biological Chemistry* **2005**, 280, (22), 21220-21230.
- 95 Hu, X., Kohler, K., Falick, A. M., Moore, A. M., Jones, P. R. & Vierra, C. Spider egg case core fibers: trimeric complexes assembled from TuSp1, ECP-1, and ECP-2. *Biochemistry* **2006**, 45, (11), 3506-3516.
- 96 Foelix, R. F. *Biology of spiders*. 2nd ed.; Oxford University Press: Oxford, UK, **1996**.
- 97 Vasanthavada, K., Hu, X., Falick, A. M., La Mattina, C., Moore, A. M. F., Jones, P. R., Yee, R., Reza, R., Tuton, T. & Vierra, C. Aciniform spidroin, a constituent of egg case sacs and wrapping silk fibers from the black widow spider *Latrodectus hesperus*. *Journal of Biological Chemistry* **2007**, 282, (48), 35088-35097.
- 98 Hayashi, C. Y., Blackledge, T. A. & Lewis, R. V. Molecular and mechanical characterization of aciniform silk: Uniformity of iterated sequence modules in a novel member of the spider silk fibroin gene family. *Molecular Biology and Evolution* **2004**, 21, (10), 1950-1959.
- 99 Blackledge, T. A. & Hayashi, C. Y. Silken toolkits: biomechanics of silk fibers spun by the orb web spider *Argiope argentata* (Fabricius 1775). *Journal of Experimental Biology* **2006**, 209, (13), 2452-2461.
- 100 Spöner, A., Vater, W., Monajembashi, S., Unger, E., Grosse, F. & Weisshart, K. Composition and hierarchical organisation of a spider silk. *PLoS ONE* **2007**, 2, (10), e998.
- 101 Hinman, M. B. & Lewis, R. V. Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber. *Journal of Biological Chemistry* **1992**, 267, (27), 19320-19324.
- 102 Li, S. F. Y., Mcghee, A. J. & Tang, S. L. New Internal Structure of Spider Dragline Silk Revealed by Atomic-Force Microscopy. *Biophysical Journal* **1994**, 66, (4), 1209-1212.
- 103 Frische, S., Maunsbach, A. B. & Vollrath, F. Elongate cavities and skin-core structure in *Nephila* spider silk observed by electron microscopy. *Journal of Microscopy* **1998**, 189, 64-70.
- 104 Schulz, S. Composition of the silk lipids of the spider *Nephila clavipes*. *Lipids* **2001**, 36, (6), 637-647.
- 105 Liu, Y., Shao, Z. Z. & Vollrath, F. Relationships between supercontraction and mechanical properties of spider silk. *Nature Materials* **2005**, 4, (12), 901-905.
- 106 Weiskopf, A., Senecal, K., Vouros, P., Kaplan, D. & Mello, C. M. The carbohydrate composition of a spider silk: *Nephila clavipes* Dragline. *Glycobiology* **1996**, 6, (7), 1703-1703.
- 107 Heidebrecht, A. & Scheibel, T. Recombinant production of spider silk proteins. *Advances in Applied Microbiology* **2013**, 82, 115-153.
- 108 O'Brien, J. P., Fahnestock, S. R., Termonia, Y. & Gardner, K. C. H. Nylons from nature: Synthetic analogs to spider silk. *Advanced Materials* **1998**, 10, (15), 1185-1195.

- 109 Knight, D. P., Knight, M. M. & Vollrath, F. Beta transition and stress-induced phase separation in the spinning of spider dragline silk. *International Journal of Biological Macromolecules* **2000**, 27, (3), 205-210.
- 110 Sponner, A., Unger, E., Grosse, F. & Klaus, W. Differential polymerization of the two main protein components of dragline silk during fibre spinning. *Nature Materials* **2005**, 4, (10), 772-775.
- 111 Craig, C. L. & Riekel, C. Comparative architecture of silks, fibrous proteins and their encoding genes in insects and spiders. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology* **2002**, 133, (4), 493-507.
- 112 Thiel, B. L. & Viney, C. Spider major ampullate silk (drag line): Smart composite processing based on imperfect crystals. *Journal of Microscopy-Oxford* **1997**, 185, 179-187.
- 113 Thiel, B. L., Guess, K. B. & Viney, C. Non-periodic lattice crystals in the hierarchical microstructure of spider (major ampullate) silk. *Biopolymers* **1997**, 41, (7), 703-719.
- 114 van Beek, J. D., Hess, S., Vollrath, F. & Meier, B. H. The molecular structure of spider dragline silk: folding and orientation of the protein backbone. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, 99, (16), 10266-10271.
- 115 Termonia, Y. Molecular Modeling of Spider Silk Elasticity. *Macromolecules* **1994**, 27, (25), 7378-7381.
- 116 Crews, S. C. & Opell, B. D. The features of capture threads and orb-webs produced by unfed *Cyclosa turbinata* (Araneae : Araneidae). *Journal of Arachnology* **2006**, 34, (2), 427-434.
- 117 Tillinghast, E. K. & Townley, M. A. Silk Glands of Araneid Spiders - Selected Morphological and Physiological-Aspects. *Silk Polymers* **1994**, 544, 29-44.
- 118 Breslauer, D. N., Lee, L. P. & Müller, S. J. Simulation of flow in the silk gland. *Biomacromolecules* **2009**, 10, (1), 49-57.
- 119 Bell, A. L. & Peakall, D. B. Changes in fine structure during silk protein production in the ampullate gland of the spider *Araneus sericatus*. *The Journal of Cell Biology* **1969**, 42, (1), 284-295.
- 120 Kovoer, J. & Zylberbe.L. Morphology and Ultrastructure of Ampullate Gland Duct in *Araneus-Diadematus* Clerck (Arachnida, Araneidae). *Zeitschrift Fur Zellforschung Und Mikroskopische Anatomie* **1972**, 128, (2), 188-211.
- 121 Plazaola, A. & Candelas, G. C. Stimulation of Fibroin Synthesis Elicits Ultrastructural Modifications in Spider Silk Secretory-Cells. *Tissue & Cell* **1991**, 23, (2), 277-284.
- 122 Hijirida, D. H., Do, K. G., Michal, C., Wong, S., Zax, D. & Jelinski, L. W. <sup>13</sup>C NMR of *Nephila clavipes* major ampullate silk gland. *Biophysical Journal* **1996**, 71, (6), 3442-3447.
- 123 Knight, D. P. & Vollrath, F. Changes in element composition along the spinning duct in a *Nephila* spider. *Naturwissenschaften* **2001**, 88, (4), 179-182.
- 124 Jin, H. J. & Kaplan, D. L. Mechanism of silk processing in insects and spiders. *Nature* **2003**, 424, (6952), 1057-1061.
- 125 Eisoldt, L., Hardy, J. G., Heim, M. & Scheibel, T. R. The role of salt and shear on the storage and assembly of spider silk proteins. *Journal of Structural Biology* **2010**, 170, (2), 413-419.
- 126 Exler, J. H., Hummerich, D. & Scheibel, T. The amphiphilic properties of spider silks are important for spinning. *Angewandte Chemie-International Edition* **2007**, 46, (19), 3559-3562.
- 127 Sponner, A., Vater, W., Rommerskirch, W., Vollrath, F., Unger, E., Grosse, F. & Weisshart, K. The conserved C-termini contribute to the properties of spider silk

- fibroins. *Biochemical and Biophysical Research Communications* **2005**, 338, (2), 897-902.
- 128 Lu, Q., Zhu, H., Zhang, C., Zhang, F., Zhang, B. & Kaplan, D. L. Silk self-assembly mechanisms and control – From thermodynamics to kinetics. *Biomacromolecules* **2012**, 13, (3), 826-832.
- 129 Vollrath, F. & Knight, D. P. Structure and function of the silk production pathway in the spider *Nephila edulis*. *International Journal of Biological Macromolecules* **1999**, 24, (2-3), 243-249.
- 130 Chen, X., Knight, D. P. & Vollrath, F. Rheological characterization of *Nephila* spidroin solution. *Biomacromolecules* **2002**, 3, (4), 644-648.
- 131 Andersson, M., Chen, G. F., Otikovs, M., Landreh, M., Nordling, K., Kronqvist, N., Westermarck, P., Jornvall, H., Knight, S., Ridderstrale, Y., Holm, L., Meng, Q., Jaudzems, K., Chesler, M., Johansson, J. & Rising, A. Carbonic Anhydrase Generates CO<sub>2</sub> and H<sup>+</sup> That Drive Spider Silk Formation Via Opposite Effects on the Terminal Domains. *Plos Biology* **2014**, 12, (8), e1001921.
- 132 Dicko, C., Kenney, J. M., Knight, D. & Vollrath, F. Transition to a beta-sheet-rich structure in spidroin in vitro: The effects of pH and cations. *Biochemistry* **2004**, 43, (44), 14080-14087.
- 133 Vollrath, F., Knight, D. P. & Hu, X. W. Silk production in a spider involves acid bath treatment. *Proceedings of the Royal Society of London Series B-Biological Sciences* **1998**, 265, (1398), 817-820.
- 134 Foo, C. W. P., Bini, E., Hensman, J., Knight, D. P., Lewis, R. V. & Kaplan, D. L. Role of pH and charge on silk protein assembly in insects and spiders. *Applied Physics a-Materials Science & Processing* **2006**, 82, (2), 223-233.
- 135 Beckwitt, R. & Arcidiacono, S. Sequence conservation in the C-terminal region of spider silk proteins (spidroin) from *Nephila clavipes* (Tetragnathidae) and *Araneus bicentenarius* (Araneidae). *Journal of Biological Chemistry* **1994**, 269, (9), 6661-6663.
- 136 Ayoub, N. A., Garb, J. E., Tinghitella, R. M., Collin, M. A. & Hayashi, C. Y. Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. *PLoS ONE* **2007**, 2, (6), e514.
- 137 Sponner, A., Schlott, B., Vollrath, F., Unger, E., Grosse, F. & Weisshart, K. Characterization of the protein components of *Nephila clavipes* dragline silk. *Biochemistry* **2005**, 44, (12), 4727-4736.
- 138 Hagn, F., Eisoldt, L., Hardy, J. G., Vendrely, C., Coles, M., Scheibel, T. & Kessler, H. A conserved spider silk domain acts as a molecular switch that controls fibre assembly. *Nature* **2010**, 465, (7295), 239-242.
- 139 Askarieh, G., Hedhammar, M., Nordling, K., Saenz, A., Casals, C., Rising, A., Johansson, J. & Knight, S. D. Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature* **2010**, 465, (7295), 236-238.
- 140 Hagn, F., Thamm, C., Scheibel, T. & Kessler, H. pH-dependent dimerization and salt-dependent stabilization of the N-terminal domain of spider dragline silk – implications for fiber formation. *Angewandte Chemie International Edition* **2010**.
- 141 Schrodinger, L. The PyMOL Molecular Graphics System, Version 1.8. In **2015**.
- 142 Lewis, R. V., Hinman, M., Kothakota, S. & Fournier, M. J. Expression and purification of a spider silk protein: A new strategy for producing repetitive proteins. *Protein Expression and Purification* **1996**, 7, (4), 400-406.
- 143 Fahnstock, S. R. & Bedzyk, L. A. Production of synthetic spider dragline silk protein in *Pichia pastoris*. *Applied Microbiology and Biotechnology* **1997**, 47, (1), 33-39.

- 144 Wen, H. X., Lan, X. Q., Zhang, Y. S., Zhao, T. F., Wang, Y. J., Kajiura, Z. & Nakagaki, M. Transgenic silkworms (*Bombyx mori*) produce recombinant spider dragline silk in cocoons. *Molecular Biology Reports* **2010**, 37, (4), 1815-1821.
- 145 Miao, Y. G., Zhang, Y. S., Nakagaki, K., Zhao, T. F., Zhao, A. C., Meng, Y., Nakagaki, M., Park, E. Y. & Maenaka, K. Expression of spider flagelliform silk protein in *Bombyx mori* cell line by a novel Bac-to-Bac/BmNPV baculovirus expression system. *Applied Microbiology and Biotechnology* **2006**, 71, (2), 192-199.
- 146 Zhang, Y., Hu, J., Miao, Y., Zhao, A., Zhao, T., Wu, D., Liang, L., Miikura, A., Shiomi, K., Kajiura, Z. & Nakagaki, M. Expression of EGFP-spider dragline silk fusion protein in BmN cells and larvae of silkworm showed the solubility is primary limit for dragline proteins yield. *Molecular Biology Reports* **2008**, 35, (3), 329-335.
- 147 Lazaris, A., Arcidiacono, S., Huang, Y., Zhou, J. F., Duguay, F., Chretien, N., Welsh, E. A., Soares, J. W. & Karatzas, C. N. Spider silk fibers spun from soluble recombinant silk produced in mammalian cells. *Science* **2002**, 295, (5554), 472-476.
- 148 Scheller, J., Guhrs, K. H., Grosse, F. & Conrad, U. Production of spider silk proteins in tobacco and potato. *Nature Biotechnology* **2001**, 19, (6), 573-577.
- 149 Barr, L. A., Fahnestock, S. R. & Yang, J. J. Production and purification of recombinant DP1B silk-like protein in plants. *Molecular Breeding* **2004**, 13, (4), 345-356.
- 150 Karatzas, C. N., Turner, J. D. & Karatzas, A. L. Production of Biofilaments in Transgenic Animals. **1999**.
- 151 Xu, H. T., Fan, B. L., Yu, S. Y., Huang, Y. H., Zhao, Z. H., Lian, Z. X., Dai, Y. P., Wang, L. L., Liu, Z. L., Fei, J. & Li, N. Construct synthetic gene encoding artificial spider dragline silk protein and its expression in milk of transgenic mice. *Animal Biotechnology* **2007**, 18, (1), 1-12.
- 152 Williams, D. Sows' ears, silk purses and goats' milk: new production methods and medical applications for silk. *Medical device technology* **2003**, 14, (5), 9-11.
- 153 Ittah, S., Cohen, S., Garty, S., Cohn, D. & Gat, U. An essential role for the C-terminal domain of a dragline spider silk protein in directing fiber formation. *Biomacromolecules* **2006**, 7, (6), 1790-1795.
- 154 Tucker, C. L., Jones, J. A., Brinshurst, H. N., Copeland, C. G., Addison, J. B., Weber, W. S., Mou, Q. S., Yarger, J. L. & Lewis, R. V. Mechanical and Physical Properties of Recombinant Spider Silk Films Using Organic and Aqueous Solvents. *Biomacromolecules* **2014**, 15, (8), 3158-3170.
- 155 Perez-Rigueiro, J., Elices, M., Guinea, G. V., Plaza, G. R., Karatzas, C., Riekkel, C., Agullo-Rueda, F. & Daza, R. Bioinspired fibers follow the track of natural spider silk. *Macromolecules* **2011**, 44, (5), 1166-1176.
- 156 Tokareva, O., Michalczechen-Lacerda, V. A., Rech, E. L. & Kaplan, D. L. Recombinant DNA production of spider silk proteins. *Microbial Biotechnology* **2013**, 6, (6), 651-663.
- 157 Arcidiacono, S., Mello, C., Kaplan, D., Cheley, S. & Bayley, H. Purification and characterization of recombinant spider silk expressed in *Escherichia coli*. *Applied Microbiology and Biotechnology* **1998**, 49, (1), 31-38.
- 158 Rabotyagova, O. S., Cebe, P. & Kaplan, D. L. Protein-based block copolymers. *Biomacromolecules* **2011**.
- 159 Xia, X. X., Qian, Z. G., Ki, C. S., Park, Y. H., Kaplan, D. L. & Lee, S. Y. Native-sized recombinant spider silk protein produced in metabolically engineered *Escherichia coli* results in a strong fiber. *Proceedings of the National Academy of Sciences of the United States of America* **2010**, 107, (32), 14059-14063.
- 160 Huemmerich, D., Helsen, C. W., Quedzuweit, S., Oschmann, J., Rudolph, R. & Scheibel, T. Primary structure elements of spider dragline silks and their contribution to protein solubility. *Biochemistry* **2004**, 43, (42), 13604-13612.



- 161 Werten, M. W. T., Moers, A. P. H. A., Vong, T., Zuilhof, H., van Hest, J. C. M. & de Wolf, F. A. Biosynthesis of an Amphiphilic Silk-Like Polymer. *Biomacromolecules* **2008**.
- 162 Winkler, S. & Kaplan, D. L. Molecular biology of spider silk. *Journal of Biotechnology* **2000**, 74, (2), 85-93.
- 163 Hayashi, C. Y., Shipley, N. H. & Lewis, R. V. Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. *International Journal of Biological Macromolecules* **1999**, 24, (2-3), 271-275.
- 164 Scheibel, T. Spider silks: recombinant synthesis, assembly, spinning, and engineering of synthetic proteins. *Microbial Cell Factories* **2004**, 3, (1), 14.
- 165 Morris, A. L., Macarthur, M. W., Hutchinson, E. G. & Thornton, J. M. Stereochemical Quality of Protein-Structure Coordinates. *Proteins-Structure Function and Genetics* **1992**, 12, (4), 345-364.
- 166 Rising, A., Nimmervoll, H., Grip, S., Fernandez-Arias, A., Storckenfeldt, E., Knight, D. P., Vollrath, F. & Engstrom, W. Spider silk proteins - Mechanical property and gene sequence. *Zoological Science* **2005**, 22, (3), 273-281.
- 167 Parkhe, A. D., Seeley, S. K., Gardner, K., Thompson, L. & Lewis, R. V. Structural studies of spider silk proteins in the fiber. *Journal of Molecular Recognition* **1997**, 10, (1), 1-6.
- 168 Bram, A., Branden, C. I., Craig, C., Snigireva, I. & Riek, C. X-ray diffraction from single fibres of spider silk. *Journal of Applied Crystallography* **1997**, 30, 390-392.
- 169 Simmons, A., Ray, E. & Jelinski, L. W. Solid-State C-13 Nmr of *Nephila clavipes* dragline silk establishes structure and identity of crystalline regions *Macromolecules* **1994**, 27, (18), 5235-5237.
- 170 Kummerlen, J., vanBeek, J. D., Vollrath, F. & Meier, B. H. Local structure in spider dragline silk investigated by two-dimensional spin-diffusion nuclear magnetic resonance. *Macromolecules* **1996**, 29, (8), 2920-2928.
- 171 Thiel, B. L. & Viney, C. beta sheets and spider silk. *Science* **1996**, 273, (5281), 1480-1481.
- 172 Work, R. W. Dimensions, birefringences, and force-elongation behavior of Major and Minor Ampullate silk fibers from orb-web spinning spiders – Effects of wetting on these properties. *Textile Research Journal* **1977**, 47, (10), 650-662.
- 173 Heidebrecht, A., Eisoldt, L., Diehl, J., Schmidt, A., Geffers, M., Lang, G. & Scheibel, T. Biomimetic Fibers Made of Recombinant Spidroins with the Same Toughness as Natural Spider Silk. *Advanced Materials* **2015**, 27, (13), 2189-2194.
- 174 Dong, Z., Lewis, R. V. & Middaugh, C. R. Molecular mechanism of spider silk elasticity. *Archives of Biochemistry and Biophysics* **1991**, 284, (1), 53-57.
- 175 Gray, G. M., van der Vaart, A., Guo, C. C., Jones, J., Onofrei, D., Cherry, B. R., Lewis, R. V., Yarger, J. L. & Holland, G. P. Secondary Structure Adopted by the Gly-Gly-X Repetitive Regions of Dragline Spider Silk. *International Journal of Molecular Sciences* **2016**, 17, (12), 2023-2037.
- 176 Lefevre, T., Rousseau, M. E. & Pezolet, M. Protein secondary structure and orientation in silk as revealed by Raman spectromicroscopy. *Biophysical Journal* **2007**, 92, (8), 2885-2895.
- 177 Motriuk-Smith, D., Smith, A., Hayashi, C. Y. & Lewis, R. V. Analysis of the conserved N-terminal domains in major ampullate spider silk proteins. *Biomacromolecules* **2005**, 6, (6), 3152-3159.
- 178 Rising, A., Hjalm, G., Engstrom, W. & Johansson, J. N-terminal nonrepetitive domain common to dragline, flagelliform, and cylindriform spider silk proteins. *Biomacromolecules* **2006**, 7, (11), 3120-3124.
- 179 Garb, J. E., Ayoub, N. A. & Hayashi, C. Y. Untangling spider silk evolution with spidroin terminal domains. *BMC Evolutionary Biology* **2010**, 10, 243.

- 180 Otikovs, M., Chen, G. F., Nordling, K., Landreh, M., Meng, Q., Jornvall, H., Kronqvist, N., Rising, A., Johansson, J. & Jaudzems, K. Diversified Structural Basis of a Conserved Molecular Mechanism for pH-Dependent Dimerization in Spider Silk N-Terminal Domains. *Chembiochem* **2015**, 16, (12), 1720-1724.
- 181 Eisoldt, L., Thamm, C. & Scheibel, T. Review the role of terminal domains during storage and assembly of spider silk proteins. *Biopolymers* **2012**, 97, (6), 355-361.
- 182 Nielsen, H., Engelbrecht, J., Brunak, S. & vonHeijne, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Engineering* **1997**, 10, (1), 1-6.
- 183 Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. & Drummond, A. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, 28, (12), 1647-1649.
- 184 Drozdetskiy, A., Cole, C., Procter, J. & Barton, G. J. JPred4: a protein secondary structure prediction server. *Nucleic Acids Research* **2015**, 43, (W1), W389-W394.
- 185 Jaudzems, K., Askarieh, G., Landreh, M., Nordling, K., Hedhammar, M., Jornvall, H., Rising, A., Knight, S. D. & Johansson, J. pH-dependent dimerization of spider silk N-terminal domain requires relocation of a wedged tryptophan side chain. *Journal of Molecular Biology* **2012**, 422, (4), 477-487.
- 186 Ries, J., Schwarze, S., Johnson, C. M. & Neuweiler, H. Microsecond Folding and Domain Motions of a Spider Silk Protein Structural Switch. *Journal of the American Chemical Society* **2014**, 136, (49), 17136-17144.
- 187 Landreh, M., Askarieh, G., Nordling, K., Hedhammar, M., Rising, A., Casals, C., Astorga-Wells, J., Alvelius, G., Knight, S. D., Johansson, J., Jörnvall, H. & Bergman, T. A pH-dependent dimer lock in spider silk protein. *Journal of Molecular Biology* **2010**, 404, (2), 328-336.
- 188 Gaines, W. A., Sehorn, M. G. & Marcotte, W. R. Spidroin N-terminal domain promotes a pH-dependent association of silk proteins during self-assembly. *Journal of Biological Chemistry* **2010**, 285, (52), 40745-40753.
- 189 Schwarze, S., Zwettler, F. U., Johnson, C. M. & Neuweiler, H. The N-terminal domains of spider silk proteins assemble ultrafast and protected from charge screening. *Nature Communications* **2013**, 4, (2815), 1-7.
- 190 Laity, P. R., Gilks, S. E. & Holland, C. Rheological behaviour of native silk feedstocks. *Polymer* **2015**, 67, 28-39.
- 191 Kronqvist, N., Otikovs, M., Chmyrov, V., Chen, G., Andersson, M., Nordling, K., Landreh, M., Sarr, M., Jornvall, H., Wennmalm, S., Widengren, J., Meng, Q., Rising, A., Otzen, D., Knight, S. D., Jaudzems, K. & Johansson, J. Sequential pH-driven dimerization and stabilization of the N-terminal domain enables rapid spider silk formation. *Nature Communications* **2014**, 5, (3254), 1-11.
- 192 Edsall, J. T., Martin, R. B. & Hollingworth, B. R. Ionization of Individual Groups in Dibasic Acids, with Application to the Amino and Hydroxylgroups of Tyrosine. *Proceedings of the National Academy of Sciences of the United States of America* **1958**, 44, (6), 505-518.
- 193 Nozaki, Y. & Tanford, C. Acid-Base Titrations in Concentrated Guanidine Hydrochloride. Dissociation Constants of Guanidinium Ion and of Some Amino Acids. *Journal of the American Chemical Society* **1967**, 89, (4), 736-742.
- 194 Bashford, D. Macroscopic electrostatic models for protonation states in proteins. *Frontiers in Bioscience-Landmark* **2004**, 9, 1082-1099.
- 195 Gunner, M. R., Mao, J. J., Song, Y. F. & Kim, J. Factors influencing the energetics of electron and proton transfers in proteins. What can be learned from calculations? *Biochimica Et Biophysica Acta-Bioenergetics* **2006**, 1757, (8), 942-968.

- 196 Pace, C. N., Grimsley, G. R. & Scholtz, J. M. Protein Ionizable Groups: pK Values and Their Contribution to Protein Stability and Solubility. *Journal of Biological Chemistry* **2009**, 284, (20), 13285-13289.
- 197 Wallace, J. A. & Shen, J. K. Unraveling a trap-and-trigger mechanism in the pH-sensitive self-assembly of spider silk proteins. *The Journal of Physical Chemistry Letters* **2012**, 3, (5), 658-662.
- 198 da Silva, F. L. B., Pasquali, S., Derreumaux, P. & Dias, L. G. Electrostatics analysis of the mutational and pH effects of the N-terminal domain self-association of the major ampullate spidroin. *Soft Matter* **2016**, 12, (25), 5600-5612.
- 199 Anandakrishnan, R., Aguilar, B. & Onufriev, A. V. H++3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Research* **2012**, 40, (W1), W537-W541.
- 200 Atkison, J. H., Parnham, S., Marcotte, W. R. & Olsen, S. K. Crystal Structure of the *Nephila clavipes* Major Ampullate Spidroin 1 A N-terminal Domain Reveals Plasticity at the Dimer Interface. *Journal of Biological Chemistry* **2016**, 291, (36), 19006-19017.
- 201 Gronau, G., Qin, Z. & Buehler, M. J. Effect of sodium chloride on the structure and stability of spider silk's N-terminal protein domain. *Biomaterials Science* **2013**, 1, (3), 276-284.
- 202 Kurut, A., Dicko, C. & Lund, M. Dimerization of Terminal Domains in Spiders Silk Proteins Is Controlled by Electrostatic Anisotropy and Modulated by Hydrophobic Patches. *Acs Biomaterials Science & Engineering* **2015**, 1, (6), 363-371.
- 203 Bini, E., Knight, D. P. & Kaplan, D. L. Mapping domain structures in silks from insects and spiders related to protein assembly. *Journal of Molecular Biology* **2004**, 335, (1), 27-40.
- 204 Beckwitt, R., Arcidiacono, S. & Stote, R. Evolution of repetitive proteins: spider silks from *Nephila clavipes* (Tetragnathidae) and *Araneus bicentenarius* (Araneidae). *Insect Biochemistry and Molecular Biology* **1998**, 28, (3), 121-130.
- 205 Ittah, S., Michaeli, A., Goldblum, A. & Gat, U. A model for the structure of the C-terminal domain of dragline spider silk and the role of its conserved cysteine. *Biomacromolecules* **2007**, 8, (9), 2768-2773.
- 206 Challis, R. J., Goodacre, S. L. & Hewitt, G. M. Evolution of spider silks: conservation and diversification of the C-terminus. *Insect Molecular Biology* **2006**, 15, (1), 45-56.
- 207 Wang, S. J., Huang, W. D. & Yang, D. W. Structure and Function of C-Terminal Domain of Aciniform Spidroin. *Biomacromolecules* **2014**, 15, (2), 468-477.
- 208 Sponner, A., Unger, E., Grosse, F. & Weisshart, K. Conserved C-termini of spidroins are secreted by the major ampullate glands and retained in the silk thread. *Biomacromolecules* **2004**, 5, (3), 840-845.
- 209 Gauthier, M., Leclerc, J., Lefevre, T., Gagne, S. M. & Auger, M. Effect of pH on the Structure of the Recombinant C-Terminal Domain of *Nephila clavipes* Dragline Silk Protein. *Biomacromolecules* **2014**, 15, (12), 4447-4454.
- 210 Humenik, M., Magdeburg, M. & Scheibel, T. Influence of repeat numbers on self-assembly rates of repetitive recombinant spider silk proteins. *Journal of Structural Biology* **2014**, 186, (3), 431-437.
- 211 Lefevre, T., Boudreault, S., Cloutier, C. & Pezolet, M. Diversity of molecular transformations involved in the formation of spider silks. *Journal of Molecular Biology* **2011**, 405, (1), 238-253.
- 212 Uversky, V. N. & Dunker, A. K. Understanding protein non-folding. *Biochimica et Biophysica Acta, Proteins and Proteomics* **2010**, 1804, (6), 1231-1264.

- 213 Humenik, M., Smith, A. M., Arndt, S. & Scheibel, T. Ion and seed dependent fibril assembly of a spidroin core domain. *Journal of Structural Biology* **2015**, 191, (2), 130-138.
- 214 Jahn, T. R., Makin, O. S., Morris, K. L., Marshall, K. E., Tian, P., Sikorski, P. & Serpell, L. C. The Common Architecture of Cross- $\beta$  Amyloid. *Journal of Molecular Biology* **2010**, 395, (4), 717-727.
- 215 Nguyen, H. D. & Hall, C. K. Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, 101, (46), 16180-16185.
- 216 Slotta, U., Hess, S., Spiess, K., Stromer, T., Serpell, L. & Scheibel, T. Spider silk and amyloid fibrils: A structural comparison. *Macromolecular Bioscience* **2007**, 7, (2), 183-188.
- 217 Rammensee, S., Slotta, U., Scheibel, T. & Bausch, A. R. Assembly mechanism of recombinant spider silk proteins. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, 105, (18), 6590-6595.
- 218 Vezy, C., Hermanson, K. D., Scheibel, T. & Bausch, A. R. Interfacial rheological properties of recombinant spider-silk proteins. *Biointerphases* **2009**, 4, (3), 43-46.
- 219 Rising, A., Widhe, M., Johansson, J. & Hedhammar, M. Spider silk proteins: recent advances in recombinant production, structure–function relationships and biomedical applications. *Cellular and Molecular Life Sciences* **2011**, 68, (2), 169-184.
- 220 Stark, M., Grip, S., Rising, A., Hedhammar, M., Engstrom, W., Hjalms, G. & Johansson, J. Macroscopic fibers self-assembled from recombinant miniature spider silk proteins. *Biomacromolecules* **2007**, 8, (5), 1695-1701.
- 221 Bauer, J., Schaal, D., Eisoldt, L., Schweimer, K., Schwarzinger, S. & Scheibel, T. Acidic Residues Control the Dimerization of the N-terminal Domain of Black Widow Spiders' Major Ampullate Spidroin 1. *Scientific Reports* **2016**, 6, (34442), 1-9.
- 222 Schaal, D., Bauer, J., Schweimer, K., Scheibel, T., Rosch, P. & Schwarzinger, S. Resonance assignment of an engineered amino-terminal domain of a major ampullate spider silk with neutralized charge cluster. *Biomolecular Nmr Assignments* **2016**, 10, (1), 199-202.
- 223 Bauer, J. & Scheibel, T. Dimerization of the Conserved N-Terminal Domain of a Spider Silk Protein Controls the Self-Assembly of the Repetitive Core Domain. *Biomacromolecules* **2017**, 18, (8), 2521-2528.
- 224 Bauer, J. & Scheibel, T. Conformational Stability and Interplay of Helical N- and C-Terminal Domains with Implications on Major Ampullate Spidroin Assembly. *Biomacromolecules* **2017**, 18, (3), 835-845.
- 225 Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Sayers, E. W. GenBank. *Nucleic Acids Research* **2016**, 44, (D1), D67-D72.
- 226 Prlic, A., Bliven, S., Rose, P. W., Bluhm, W. F., Bizon, C., Godzik, A. & Bourne, P. E. Pre-calculated protein structure alignments at the RCSB PDB website. *Bioinformatics* **2010**, 26, (23), 2983-2985.
- 227 Shindyalov, I. N. & Bourne, P. E. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Engineering* **1998**, 11, (9), 739-747.
- 228 Ye, Y. Z. & Godzik, A. Flexible structure alignment by chaining aligned fragment pairs allowing twists. *Bioinformatics* **2003**, 19, ii246-ii255.
- 229 Thamm, C. & Scheibel, T. Recombinant Production, Characterization, and Fiber Spinning of an Engineered Short Major Ampullate Spidroin (MaSp1s). *Biomacromolecules* **2017**, 18, (4), 1365-1372.

- 230 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic Local Alignment Search Tool. *Journal of Molecular Biology* **1990**, 215, (3), 403-410.

## 5. PUBLIKATIONSLISTE

- I. **Bauer, J.**; Schaal, D.; Eisoldt, L.; Schweimer, K.; Schwarzinger, S. & Scheibel, T. (2016) Acidic Residues Control the Dimerization of the N-terminal Domain of Black Widow Spiders' Major Ampullate Spidroin 1. *Scientific Reports* 6: 34442.
- II. Schaal, D.; **Bauer, J.**; Schweimer, K.; Scheibel, T. & Schwarzinger, S. (2016) Resonance assignment of an engineered amino-terminal domain of a major ampullate spider silk with neutralized charge cluster. *Biomolecular NMR Assignments* 1: 199.
- III. **Bauer, J.** & Scheibel, T. (2017) Dimerization of the Conserved N-Terminal Domain of a Spider Silk Protein Controls the Self-Assembly of the Repetitive Core Domain. *Biomacromolecules* 18 (8): 2521.
- IV. **Bauer, J.** & Scheibel, T. (2017) Conformational Stability and Interplay of Helical N- and C-Terminal Domains with Implications on Major Ampullate Spidroin Assembly. *Biomacromolecules* 18 (3): 835.

## 6. DARSTELLUNG DES EIGENANTEILS

Diese Dissertation besteht aus vier Publikationen, die zum Teil in Zusammenarbeit mit Kooperationspartnern erarbeitet wurden. Nachfolgend werden die einzelnen Beiträge aller Autoren zu den jeweiligen Veröffentlichungen dargestellt.

### Publikation I

**Bauer, J.;** Schaal, D.; Eisoldt, L.; Schweimer, K.; Schwarzinger, S. & Scheibel, T. (2016) Acidic Residues Control the Dimerization of the N-terminal Domain of Black Widow Spiders' Major Ampullate Spidroin 1. *Scientific Reports* 6: 34442.

Das Konzept der Studie wurde kooperativ von Daniel Schaal, Lukas Eisoldt, Stephan Schwarzinger, Thomas Scheibel und mir entworfen. Die rekombinanten Proteine wurden unter der Betreuung von Thomas Scheibel von mir hergestellt und gereinigt. Die Proteine wurden von mir Fluoreszenz- und CD-spektroskopisch sowie mit Hilfe von SEC-MALS Messungen charakterisiert. Die NMR-Spektroskopie wurde von Daniel Schaal unter der Betreuung von Kristian Schweimer und Stephan Schwarzinger durchgeführt. Die Ergebnisse wurden von mir mit allen Autoren wissenschaftlich diskutiert. Das Manuskript wurde von mir angefertigt und zusammen mit Daniel Schaal, Stephan Schwarzinger und Thomas Scheibel fertig gestellt.

### Publikation II

Schaal, D.; **Bauer, J.;** Schweimer, K.; Scheibel, T. & Schwarzinger, S. (2016) Resonance assignment of an engineered amino-terminal domain of a major ampullate spider silk with neutralized charge cluster. *Biomolecular NMR Assignments* 1: 199.

Die Konzeption des Artikels wurde in Zusammenarbeit zwischen Daniel Schaal, Kristian Schweimer, Stephan Schwarzinger, Thomas Scheibel und mir entworfen. Die Isotopenmarkierten Proteine wurden von mir produziert und von Daniel Schaal NMR-spektroskopisch analysiert. Die Resonanzzuordnung erfolgte von Daniel Schaal unter der Betreuung von Kristian Schweimer und Stephan Schwarzinger. Nach wissenschaftlicher Diskussion mit allen Autoren wurde das Manuskript von Daniel Schaal verfasst und zusammen mit Kristian Schweimer, Stephan Schwarzinger, Thomas Scheibel und mir fertig gestellt.

### **Publikation III**

**Bauer, J. & Scheibel, T. (2017)** Dimerization of the Conserved N-Terminal Domain of a Spider Silk Protein Controls the Self-Assembly of the Repetitive Core Domain. *Biomacromolecules* 18 (8): 2521.

Die Studie wurde von Thomas Scheibel und mir konzeptioniert. Die rekombinanten Proteine wurden von mir produziert, gereinigt und analysiert. Das Projekt wurde von Thomas Scheibel betreut. Das Manuskript wurde von mir erstellt und von Thomas Scheibel fertig gestellt.

### **Publikation IV**

**Bauer, J. & Scheibel, T. (2017)** Conformational Stability and Interplay of Helical N- and C-Terminal Domains with Implications on Major Ampullate Spidroin Assembly. *Biomacromolecules* 18 (3): 835.

Das Konzept der Publikation wurde von Thomas Scheibel und mir entworfen. Die native PAGE und SEC-MALS Analysen wurden von Masterstudentin Corinna Drescher im Rahmen eines Forschungspraktikums unter meiner Betreuung durchgeführt. Die restlichen Experimente wurden von mir unter der Betreuung von Thomas Scheibel durchgeführt und ausgewertet. Das Manuskript wurde von mir erstellt und von Thomas Scheibel fertig gestellt.



**TEILARBEITEN**

Teilarbeit I wurde 2016 in *Scientific Reports* veröffentlicht.

Bauer, J.; Schaal, D.; Eisoldt, L.; Schweimer, K.; Schwarzingen, S. & Scheibel, T. (2016) Acidic Residues Control the Dimerization of the N-terminal Domain of Black Widow Spiders' Major Ampullate Spidroin 1. *Scientific Reports* 6: 34442.

Im Folgenden wird die Publikation gemäß der *Creative Commons CC-BY* Lizenz (2016) reproduziert.

# SCIENTIFIC REPORTS

OPEN

## Acidic Residues Control the Dimerization of the N-terminal Domain of Black Widow Spiders' Major Ampullate Spidroin 1

Received: 21 March 2016  
Accepted: 13 September 2016  
Published: 29 September 2016

Joschka Bauer<sup>1</sup>, Daniel Schaal<sup>2,3</sup>, Lukas Eisoldt<sup>1</sup>, Kristian Schweimer<sup>2,3</sup>,  
Stephan Schwarzsinger<sup>2,3</sup> & Thomas Scheibel<sup>1,3,4,5,6</sup>

Dragline silk is the most prominent amongst spider silks and comprises two types of major ampullate spidroins (MaSp) differing in their proline content. In the natural spinning process, the conversion of soluble MaSp into a tough fiber is, amongst other factors, triggered by dimerization and conformational switching of their helical amino-terminal domains (NRN). Both processes are induced by protonation of acidic residues upon acidification along the spinning duct. Here, the structure and monomer-dimer-equilibrium of the domain NRN1 of *Latrodectus hesperus* MaSp1 and variants thereof have been investigated, and the key residues for both could be identified. Changes in ionic composition and strength within the spinning duct enable electrostatic interactions between the acidic and basic pole of two monomers which prearrange into an antiparallel dimer. Upon naturally occurring acidification this dimer is stabilized by protonation of residue E114. A conformational change is independently triggered by protonation of clustered acidic residues (D39, E76, E81). Such step-by-step mechanism allows a controlled spidroin assembly in a pH- and salt sensitive manner, preventing premature aggregation of spider silk proteins in the gland and at the same time ensuring fast and efficient dimer formation and stabilization on demand in the spinning duct.

Spider dragline silk is used by orb weavers as a lifeline to escape from predators and as a stabilizing frame in their webs due to its excellent mechanical properties including an outstanding toughness. It comprises proteins produced in the major ampullate silk gland classified as major ampullate spidroins MaSp1<sup>1</sup> and MaSp2<sup>2</sup> dependent on their amino acid composition, mainly their proline content which is low (<1%) in MaSp1 and high (~9%) in MaSp2<sup>3</sup>. Spidroins are stored in a soluble form at high concentrations (>30% w/v) in the presence of NaCl in the lumen of the silk gland<sup>4,5</sup>, and fiber assembly is initiated upon passage of the spidroin solution through the spinning duct<sup>6</sup>. The pH decreases along the duct from pH 7.2 to 6.0 or below, depending on the spider species, and sodium chloride is replaced by potassium phosphate, both of which together with shear forces induce spidroin assembly<sup>5,7–11</sup>. Within the spinning dope, the large repetitive MaSp core domains (>1000 amino acids) are assumed to be intrinsically disordered, while the short (~100–150 amino acids) non-repetitive N- (NRN) and C-terminal domains (NRC) exhibit well defined  $\alpha$ -helical structures<sup>3,12,13</sup>. These terminal domains control spidroin solubility in the dope and initiate fiber assembly upon passage through the spinning duct. Fiber assembly is accompanied by conversion of the intrinsically disordered core domain into a  $\beta$ -sheet-rich (~25–40%) one<sup>14–19</sup>. Due to their important function during assembly, the sequences and therefore the three-dimensional fold of NRC and especially NRN are highly conserved, not only of individual silks between spider species, but also between different silk types within one species<sup>20–24</sup>. While NRC of most MaSp are permanent disulphide-linked dimers even under storage conditions in the gland, dimerization of NRN is triggered in a pH and salt dependent

<sup>1</sup>Lehrstuhl Biomaterialien, Universität Bayreuth, Fakultät für Ingenieurwissenschaften, Bayreuth, Germany.

<sup>2</sup>Lehrstuhl Biopolymere, Universität Bayreuth, Bayrisches Geoinstitut, Bayreuth, Germany. <sup>3</sup>Forschungszentrum für Bio-Makromoleküle (BIOmac), Universität Bayreuth, Bayreuth, Germany. <sup>4</sup>Bayreuther Zentrum für Kolloide und Grenzflächen (BZKG), Universität Bayreuth, Bayreuth, Germany. <sup>5</sup>Bayreuther Materialzentrum (BayMat), Universität Bayreuth, Bayreuth, Germany. <sup>6</sup>Bayreuther Zentrum für Molekulare Biowissenschaften (BZMB), Universität Bayreuth, Bayreuth, Germany. Correspondence and requests for materials should be addressed to T.S. (email: thomas.scheibel@bm.uni-bayreuth.de)

manner within the duct<sup>25–27</sup>. The previously determined structure of *Euprosthenops australis* (NRN1<sub>E.a.</sub>) as well as of *Latrodectus hesperus* (NRN1<sub>L.h.</sub>) MaSp1 revealed a dipolar charge distribution, with charged amino acids being grouped into a basic NH<sub>2</sub><sup>−</sup> and an acidic COOH-terminal pole<sup>26,28,29</sup>. In NRN1<sub>E.a.</sub>, an acidic cluster, comprising at least six surface exposed aspartic and glutamic acid residues, has been assumed to mediate a pH-dependent dimerization upon conversion from a dimer-incompatible into a dimer-favouring conformation<sup>28</sup>. The monomeric wildtype (wt) structure, denoted as conformation I, is stabilized both in the presence of sodium chloride and at a neutral pH (7.2). Acidification triggers the formation of conformation II upon dimerization. Previously, the conformational change was monitored by a bathochromic shift of the fluorescence emission of a single, naturally occurring Trp residue (W9) indicating its relocation into a more hydrophilic environment on the surface<sup>12,28</sup>. Underlining their likely importance for this conformational change, one aspartic acid (D39) and three glutamic acid residues (E76, E81, E114) are the most conserved residues within NRN among all spider species so far investigated<sup>20–24</sup>.

Here, the dimerization and conformational changes of the amino-terminal domain of *Latrodectus hesperus* MaSp1 (NRN1<sub>L.h.</sub>) have been investigated in more detail. A variety of different techniques such as tryptophan fluorescence, near- and far-UV circular dichroism, multi-angle light scattering as well as two-dimensional NMR-spectroscopy were applied to characterize wtNRN1<sub>L.h.</sub> and variants thereof. The role of the most conserved acidic residues was identified by replacing them by their non-titratable analogues asparagine and glutamine to mimic the protonated state caused by acidification within the spinning duct. While protonation of the single side chain E114 upon acidification is sufficient to trigger dimer formation, the structural conversion into the final dimer conformation is mediated by the acidic cluster. Charged residues D39 and E81 prearrange the antiparallel dimer by electrostatic interactions with basic residues of the second subunit. Dimerization and structural rearrangement occur independently and are controlled by separated regions of the domain.

## Materials and Methods

**Protein production.** Genes encoding the amino-terminal domain of *Latrodectus hesperus* major ampullate spidroin 1 were cloned into pET28a, containing a sequence encoding a His6-SUMO-tag (Novagen). Mutations were introduced using a QuikChange<sup>®</sup> Site-Directed Mutagenesis Kit (Agilent Technologies). Genes were expressed in *Escherichia coli* strain BL21 (DE3) (Agilent Technologies). The cells were grown to an OD<sub>600</sub> = 8 in ZYM-5052 autoinduction medium<sup>30</sup> containing kanamycin for 17 h at 30 °C. Proteins were purified by nickel-NTA chromatography (HisTrap FF, GE Healthcare) and size exclusion chromatography (HiLoad<sup>™</sup> 26/60 Superdex<sup>™</sup> 75 pg, GE Healthcare). The His6-SUMO-tag was cleaved off by addition of a SUMO-protease and incubation for 1 h at RT. SUMO-tag and protease were separated from the spidroin using a second nickel-NTA chromatography step.

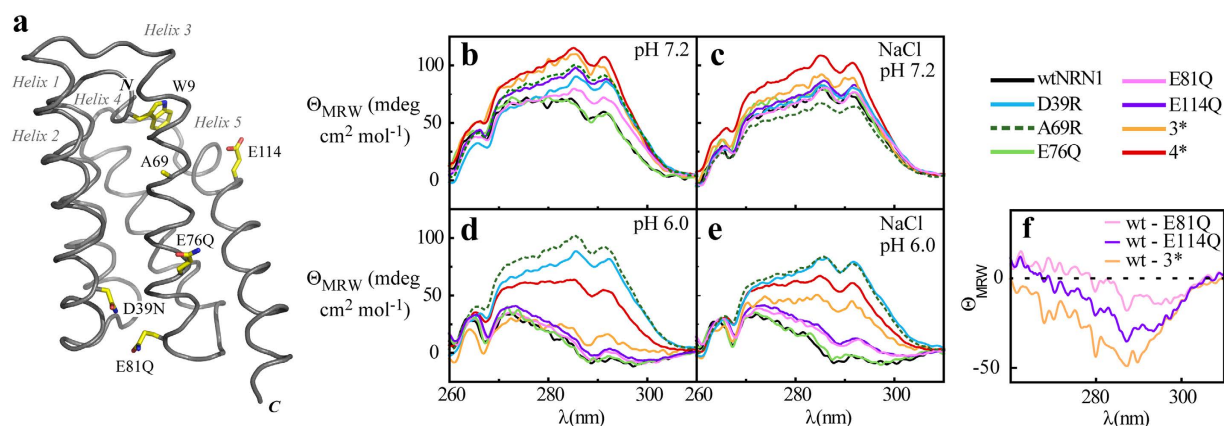
**Fluorescence spectroscopy.** Tryptophan fluorescence was analysed using a Jasco FP-6500 spectrofluorometer with a 3 mm path length. The spectra were recorded at an excitation wavelength of 295 nm and an emission wavelength between 300 and 450 nm. The pH-titration was performed by successively adding 5 µl of 0.1 M H<sub>3</sub>PO<sub>4</sub> to 1 ml of protein solution (14.2 µM).

**CD spectroscopy.** Circular dichroism spectroscopy measurements were performed using a Jasco J-715 spectropolarimeter. CD spectra were acquired at a response time of 1 s, a scanning speed of 50 nm min<sup>−1</sup> and a bandwidth of 1 nm. Far-UV CD spectra and thermal transitions were measured at 14.2 µM using a path length of 0.1 cm. Thermal transitions were recorded at a wavelength of 222 nm and a heating rate of 60 °C h<sup>−1</sup>. Near-UV CD spectra were monitored at protein concentrations of 35.5 µM and 142.5 µM (Supplementary Fig. S1) using a path length of 0.5 cm, and, respectively, of 1.0 cm. Near-UV CD spectra were processed by applying a Savitzky-Golay filter<sup>31</sup>.

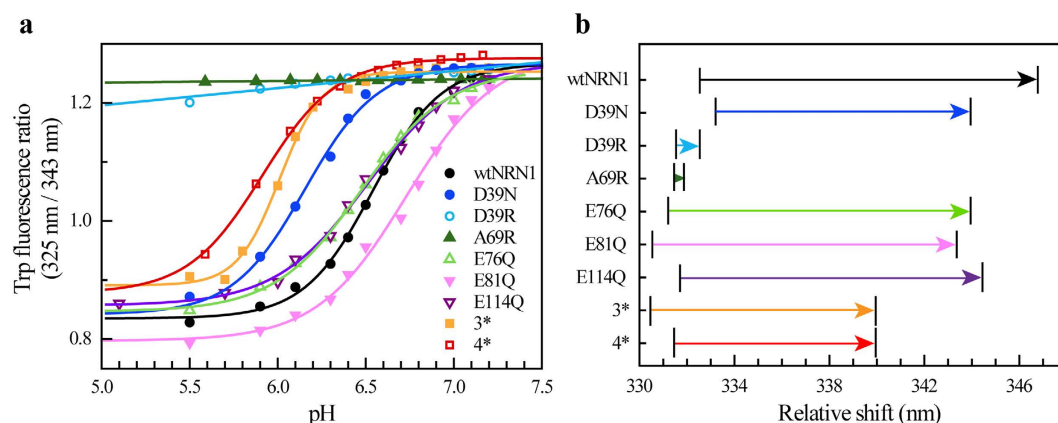
**SEC-MALS.** Size-exclusion chromatography (SEC, Superdex<sup>™</sup> 200 10/300 GL, GE Healthcare) was performed on an Agilent 1100 system to pre-separate the protein solution for the subsequent analysis. The flow rate of 20 mM sodium phosphate buffer (pH 7.2 or 6.0 ± 300 mM NaCl) was set to 0.7 ml min<sup>−1</sup>, and 250 µl of protein solution (71.2 µM) were injected at 0.2 ml min<sup>−1</sup>. Multi-angle light scattering (MALS) and quasi-elastic light scattering (QELS, WYATT) were employed to determine the molecular weight (MW) and, thus, the monomer-dimer equilibrium of NRN. The data were evaluated using the ASTRA software (WYATT).

**NMR analysis.** Isotopically enriched proteins were produced in the *E. coli* strain BL21 (DE3) grown in M9 minimal medium supplemented with <sup>15</sup>N ammonium sulphate for <sup>15</sup>N HSQC and additionally with <sup>13</sup>C glucose for triple resonance experiments. The cells were grown to an OD<sub>600</sub> = 4 at 37 °C, and genes were expressed upon addition of 1 mM IPTG for 4 h. All variants were purified in the same manner as their unlabelled analogues. 10% (v/v) D<sub>2</sub>O was added to NMR samples in 22 mM sodium phosphate, and the pH was carefully adjusted to either 7.2 or 6.0. The pH-induced conformational change of triple mutant 3\*, in which the acidic cluster was neutralized (D39N-E76Q-E81Q), shifted from pH 6.5 in wtNRN1 to 6.0 as seen in fluorescence titration experiments. Therefore, the respective HSQC experiment was recorded at pH 5.5 instead of 6.0 to ensure complete structural conversion.

All NMR spectra were recorded on a Bruker Avance 700 MHz NMR spectrometer equipped with a 5 mm TCI cryogenic and a TXI probe with Z-axis gradients, respectively. The <sup>15</sup>N HSQC experiments were recorded according to Mori *et al.*<sup>32</sup>. T<sub>1</sub> and T<sub>2</sub> <sup>15</sup>N-relaxation data were recorded for the determination of the oligomerization state of E114Q at pH 7.2. The intensities of the HSQC-type spectra were fitted to mono-exponential decays using the program curve fit (Palmer, Dept. of Biochemistry and Molecular Biophysics, Columbia University, USA). The rotational correlation time τ<sub>c</sub> was estimated using the equation



**Figure 1. Structural analysis of wildtype and variants of NRN1<sub>Lh</sub>.** (a) The lowest energy structure of the NRN1<sub>Lh</sub> variant 3\* was solved by NMR spectroscopy and is used as a structural template to show the position of the individually mutated amino acid residues (PDB accession code: 2N3E). (b–e) Near-UV CD spectra of wtNRN1<sub>Lh</sub> and variants thereof at (b) pH 7.2, (c) pH 7.2 in the presence of 300 mM NaCl, (d) pH 6.0, (e) pH 6.0 in the presence of 300 mM NaCl. (f) The near-UV CD spectra of E81Q, E114Q and 3\* were subtracted by the spectrum of wtNRN1<sub>Lh</sub> to illustrate the significant spectral differences at pH 7.2. All spectra were taken at 142.5 μM.



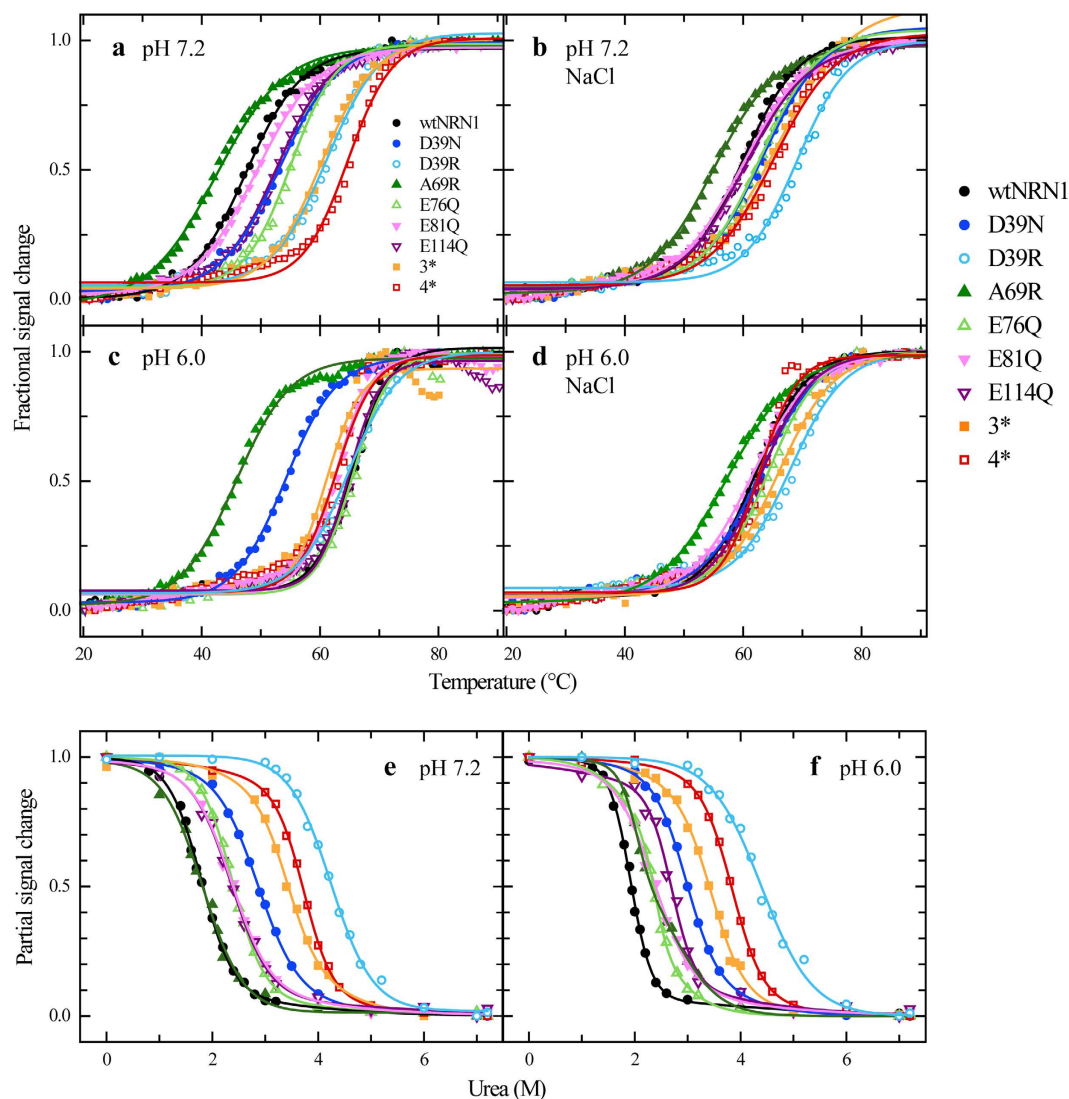
**Figure 2. Trp fluorescence analysis of NRN1<sub>Lh</sub> variants.** (a) The ratio of fluorescence at 328 nm and 343 nm indicates the proportion between conformation I and II. The lines represent fits based on a two-state model. (b) Shift of the fluorescence maximum upon lowering the pH from 7.2 to 5.5 (see also Supplementary Fig. S5).

$$\tau_c = \sqrt{\frac{6T_1}{T_2} - 7} \frac{1}{4\pi\nu_N}$$

with  $T_1$  and  $T_2$  as the longitudinal and transversal relaxation times and  $\nu_N$  as the <sup>15</sup>N nuclear frequency, respectively<sup>33</sup>. Subsequently, the molecular weight was approximated by extrapolation using a set of standard proteins<sup>34</sup>.

High resolution structural data (PDB: 2N3E) of variant 3\* were obtained by standard triple-resonance experiments and resonance assignment as described in detail in Schaal *et al.*<sup>29</sup>. The structure of 3\* is used here as a structural template to highlight the positions of the most important residues. A structural comparison to the previously published<sup>26</sup> wtNRN1<sub>Lh</sub> was not possible since the wildtype structure has not been deposited in the pdb databank (www.rcsb.org)<sup>35</sup>. Distance restraints were collected from NOESY spectra and roughly classified from very strong to very weak (2.7 to 5.5 Å) dependent on their spectral signal intensity. 120 structures were calculated using simulated annealing protocols of the XPLOR-NIH (1.2.1) software package<sup>36,37</sup>. Statistics, energetic and structural analysis were performed on the twenty lowest energy structures (Supplementary Fig. S2) using in-house software and PROCHECK<sup>38,39</sup>. Structural statistics are summarized in Supplementary Table S1.

NMR spectra were analysed and NMR distance restraints were collected using the CCPNMR software package<sup>40</sup>. HSQC images were generated by NMRViewJ (Newmoon Scientific, Westfield, NJ, USA). Images and structural alignment of 3\* were made using MacPyMOL<sup>41</sup>. The final editing of all images was performed using Adobe Illustrator CS3 (Adobe, San Jose, CA, USA).



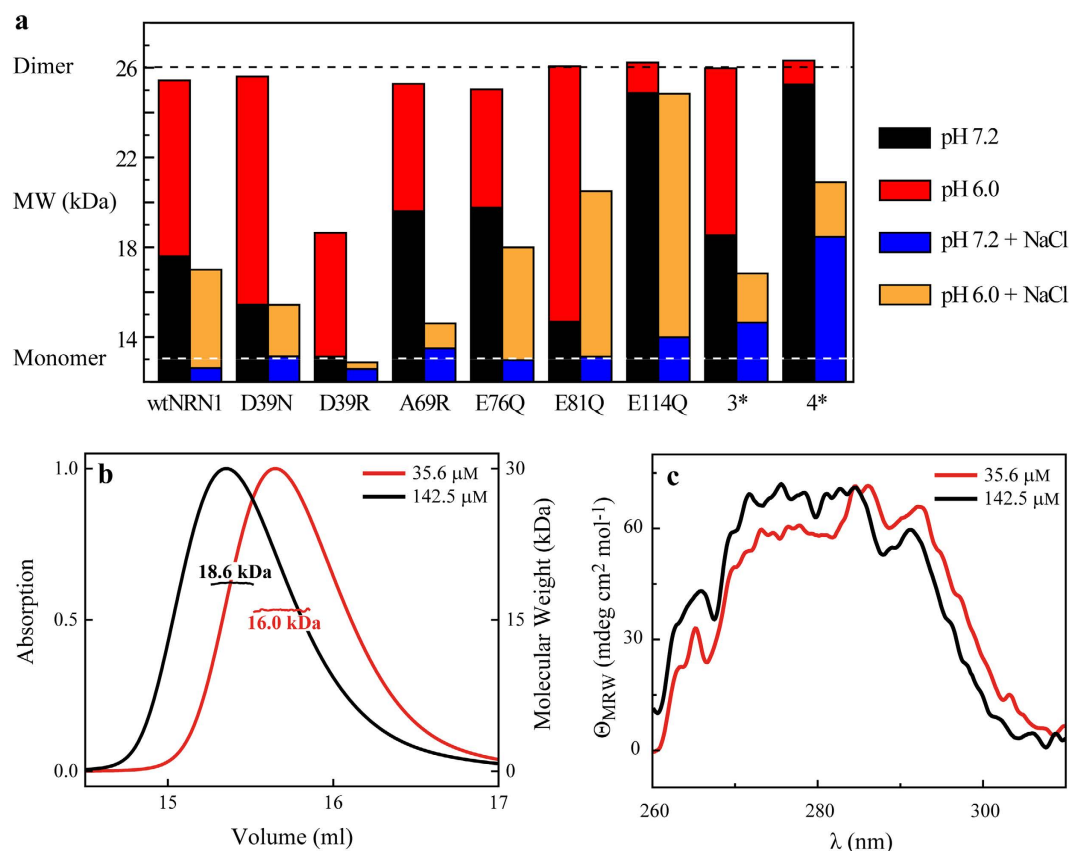
**Figure 3. Stability analysis of NRN1<sub>L,h</sub> variants.** Protein denaturation was analysed as a function of urea concentration or temperature using CD ellipticity at 222 nm. (a–d) Thermal denaturation of NRN1<sub>L,h</sub> variants at (a) pH 7.2, (b) pH 7.2 in the presence of 300 mM NaCl, (c) pH 6.0 and (d) pH 6.0 in the presence of 300 mM NaCl. (e,f) Urea titration of NRN1<sub>L,h</sub> variants at (e) pH 7.2 and (f) pH 6.0. The lines represent fits using a two-state model.

## Results

**Designing NRN1<sub>L,h</sub> variants of MaSp1 from the black widow spider with modified charge distribution.** Throughout 34 so far analysed sequences of amino-terminal domains from different types of spidroins or spider species, amino acid residues D39, E76, E81 and E114 are the most conserved ones<sup>24</sup>. These residues were replaced in the amino-terminal domain (NRN1<sub>L,h</sub>) of MaSp1 of the black widow spider (*Latrodectus hesperus*) by the non-titratable analogues asparagine and glutamine (Fig. 1a) to investigate their contribution to dimerization and conformational conversion. Four single (D39N, E76Q, E81Q, E114Q), one triple (D39N-E76Q-E81Q = 3\*) and one quadruple mutant (D39N-E76Q-E81Q-E114Q = 4\*) were investigated using several independent methods. In addition, effects of charge reversal (D39R) and charge insertion (A69R) were analysed.

**Identification of key residues of the pH- and salt-induced structural conversion.** The pH-dependent conformational state of wtNRN1<sub>L,h</sub> was determined by far- and near-UV circular dichroism (CD) spectroscopy revealing a significant pH-dependence of tertiary (Fig. 1b–e) but not of secondary structure (Supplementary Fig. S3). Mutating individually E81 or E114 clearly affected the environment of the aromatic amino acid side chains and particularly that of the single Trp (W9), and the effect was even more pronounced in the triple mutant 3\* (D39N-E76Q-E81Q) (Fig. 1b,f). Hence, negative charges appeared crucial for stabilizing the monomeric conformation I at neutral pH. Starting from wtNRN1<sub>L,h</sub> in conformation I, the five-helix bundle rearranged upon acidification into conformation II, thereby shifting W9 as indicated by a decreased ellipticity (Fig. 1b,d). Addition of





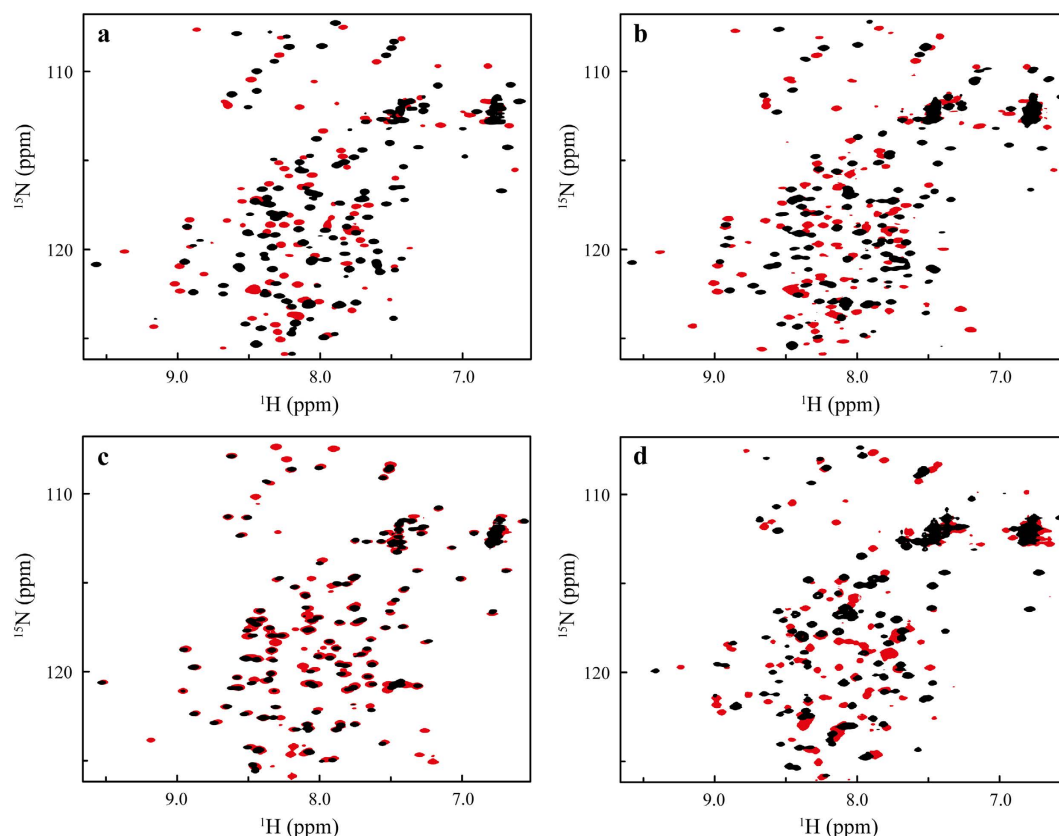
**Figure 4.** Analysis of monomer-dimer distribution of wildtype as well as variants of NRN1<sub>L<sub>h</sub></sub>. (a) Mean molecular weights indicate the ratio between monomers and dimers of NRN1<sub>L<sub>h</sub></sub> variants. MALS of 71.2 μM protein was measured at different pH values in the absence and presence of 300 mM NaCl. (b) The normalized UV signal shows the SEC profile of 35.6 μM and 142.5 μM wtNRN1<sub>L<sub>h</sub></sub> at pH 7.2. The molecular weight was determined using MALS. (c) Near-UV CD spectra of 35.6 μM and 142.5 μM wtNRN1<sub>L<sub>h</sub></sub> were taken at pH 7.2.

NaCl clearly suppressed the structural conversion of 3\* (Supplementary Fig. S4), suggesting a role of the mutated acidic residues as a salt-sensitive switch.

The importance of charge-charge interactions at the dimerization interface was emphasized by effects of charge-reversal mutations. Introduction of basic residues into the acidic cluster prevented the controlled structural conversion (Fig. 1b,d). Strikingly, the conformational change of wtNRN1<sub>L<sub>h</sub></sub> involved the relocation of residue W9 into a more hydrophilic environment, which can be seen by a fluorescence redshift with a transition point at around pH 6.5 (Fig. 2a). Neutralization of side chain D39 reduced the fluorescence redshift (Fig. 2b) and lowered the fluorescence transition to pH 6.2 (Fig. 2a). Both effects were even stronger upon deletion of the entire acidic cluster (3\*). Accordingly, the local cumulation of residues D39, E76 and E81 elevates their putative pK<sub>a</sub> to values near the physiological pH and enables wtNRN1<sub>L<sub>h</sub></sub> to completely switch between conformation I and conformation II at slightly acidic pH values (Fig. 2a). Clustering of acidic residues significantly destabilizes the wildtype protein, verified by an increased chemical and thermal stability of D39R (Fig. 3 and Supplementary Table S2). Furthermore, deletion of the acidic cluster clearly reduced the rearrangement into conformation II in the presence of salt (Fig. 1c,e and Supplementary Fig. S4). This finding suggests that the clustered residues D39, E76 and E81 cooperatively act as a pH- and salt-sensitive sensor controlling the structural conversion of NRN1<sub>L<sub>h</sub></sub>.

Introduction of an arginine residue within helix 2 (D39R) or helix 3 (A69R) clearly supported conformation I (Figs 1 and 2). Neutralizing E81 or even the entire acidic cluster (3\*) considerably altered not only the vicinity of the mutated residues, but the overall conformation I (Fig. 1a,f). Presumably, this is a consequence from lacking electrostatic repulsions between helix 2 and 3. Intramolecular electrostatic repulsions within the acidic cluster are confirmed by an increased protein stability of 3\* (Fig. 3 and Supplementary Table S2) and a simultaneously reduced conformational change in the presence of salt (Fig. 1e and Supplementary Fig. S4). The altered conformation I of 3\* does not completely prevent conformational changes upon acidification, indicating that additional so far unknown residues might be involved in attaining conformation II.

**The monomer-dimer equilibrium is not affected by structural conversion.** Monomer-dimer equilibria of the NRN1<sub>L<sub>h</sub></sub> variants were examined using multi-angle light scattering (MALS) analysis. Decreasing the pH from 7.2 to 6.0 triggered wildtype dimerization, whereas the monomeric state was stabilized in the presence of sodium chloride (Fig. 4a and Supplementary Fig. S6). Increasing the wtNRN1<sub>L<sub>h</sub></sub> concentration slightly shifted the monomer-dimer equilibrium towards the dimer state as expected (Fig. 4b). Near-UV CD spectra revealed



**Figure 5.** HSQC spectra of wtNRN1<sub>L,h</sub>, E114Q, D39R and 3\*. <sup>15</sup>N-HSQC spectra of (a) wtNRN1<sub>L,h</sub>, (b) E114Q and (c) D39R have been taken at pH 7.2 (black) and pH 6.0 (red). (d) Since the conformational change of 3\* was shifted to a lower pH, the <sup>15</sup>N-HSQC spectra of 3\* was measured at pH 7.2 (black) and at pH 5.5 (red).

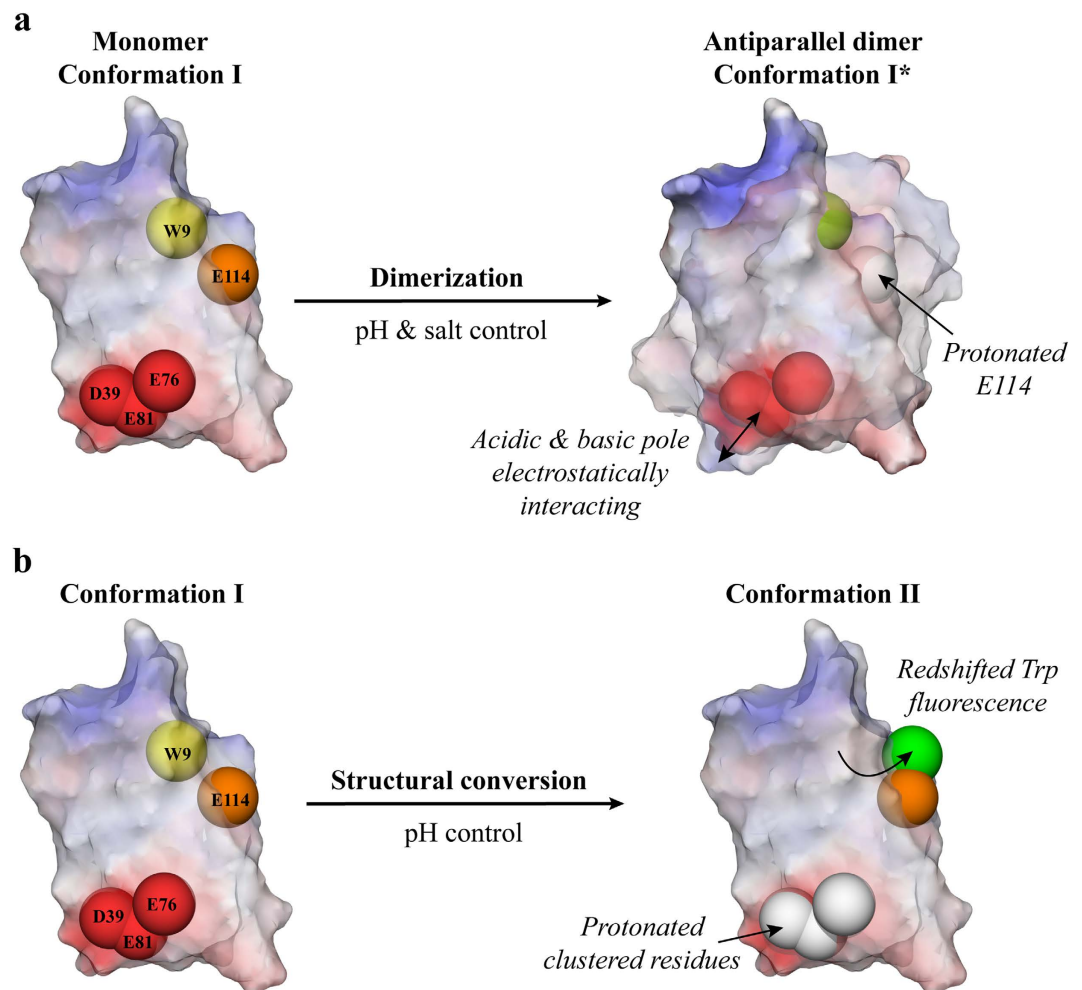
that the monomer conformation I is adjusted to a dimer-compatible conformation I\* with the Trp (W9) being slightly reoriented (Fig. 4c).

Although the dimerization interface was different due to sterical constraints by replacing a small alanine residue with arginine and its bulky side chain, the monomer-dimer equilibrium of A69R was similar to that of wtNRN1<sub>L,h</sub>. The variant A69R showed a wt-like dimerization behaviour but an inhibited structural conversion, indicating independence of both processes upon structurally rearranging NRN1<sub>L,h</sub> (Figs 1 and 2). Single mutations of residues D39 and E81 into neutral amino acids did not trigger dimerization, but even stabilized the monomeric state at neutral pH (Fig. 4a). Accordingly, the charged side chains of both residues are involved in electrostatic interactions with basic residues of the second subunit at an early stage of dimerization which is necessary to prearrange the subunits in an antiparallel manner. Contrary to the dimerization of NRN1<sub>E,a</sub><sup>10,24,28,42</sup>, structural information such as Trp fluorescence could not be used to predict the monomer-dimer equilibrium of NRN1<sub>L,h</sub>, since in this case both processes occurred independently.

Variant E114Q was considerably stabilized in the dimeric state at pH 7.2 (Fig. 4a), suggesting that protonation (i.e. neutralization) of residue E114 likely triggered and definitely stabilized dimer formation in a pH-dependent manner. Addition of NaCl-induced disassembly of E114Q dimers, and neutralization of E114 as well as the acidic cluster (variant 4\*) reduced such salt-induced dimer disassembly, indicating electrostatic interactions between the acidic and basic pole across the dimerization interface. Variant 4\* showed both a slight suppression of dimerization in the presence of NaCl as well as an unaltered conformation (Fig. 1b,c), corroborating that dimerization and structural conversion of NRN1<sub>L,h</sub> are independent and non-related events.

#### HSQC fingerprints show that changes in the acidic cluster initiate the conformational reorganization.

The independence of conformational switching and dimerization was confirmed by comparison of <sup>15</sup>N-HSQC spectra as well as <sup>15</sup>N-relaxation experiments. wtNRN1<sub>L,h</sub> showed characteristic <sup>15</sup>N-HSQC fingerprints for its distinct conformation I at pH 7.2 and conformation II at 6.0, respectively, with severe differences in the chemical shifts (Fig. 5a). In agreement with near-UV CD (Fig. 1), HSQC spectra of E114Q showed a slightly altered conformation I\* at pH 7.2 (Supplementary Fig. S7). An NMR relaxation experiment independently indicated dimerization of E114Q at neutral pH in the absence of salt by determining a rotational correlation time  $\tau_c$  of 17 ns corresponding to an estimated MW of 27 kDa (Supplementary Fig. S8). This finding provides clear evidence that in the case of NRN1<sub>L,h</sub>, residue E114 (theoretical  $pK_a$  in NRN1<sub>E,a</sub>: 4.4–6.7)<sup>43,44</sup> is the ion/pH-sensitive switch that controls dimerization. However, rearrangement into conformation II is not prevented by charge deletion of



**Figure 6.** Scheme of pH-induced dimerization and structural conversion of wtNRN1<sub>L.h.</sub> (a) Physiological solvent conditions (pH 7.2, NaCl) during spidroin storage cause deprotonated D39, E76, E81 (red spheres) and E114 (orange sphere) to stabilize monomeric NRN1<sub>L.h.</sub> in conformation I. A decreasing NaCl concentration along the spinning duct enables electrostatic interactions between the acidic and basic pole of both subunits, prearranging the dimer in an antiparallel manner. Decreasing ionic strength as well as pH effects residue E114 (orange → white sphere) which induces a slight altered conformation I\*, and, thereby, triggers stable dimer formation. (b) The structural change into the tight conformation II is initiated simultaneously in the spinning duct but is independently controlled by successive protonation of clustered D39, E76 and E81 (red → white spheres). Thereby, the only W of the domain is shifted towards the surface (yellow → green sphere).

E114 (Figs 1, 2 and 5b and Supplementary Fig. S7), showing that the structural conversion is controlled by additional residues and occurs independently from dimer formation.

Reversing the charge of residue D39 of the acidic cluster (D39R) stabilized the wt-like conformation I between pH 7.2 and pH 6.0 (Fig. 5c and Supplementary Fig. S7). Together with unchanged near-UV CD and fluorescence spectra in the same pH-range (Figs 1 and 2) this indicates that the interplay of clustered acidic residue side chains is required for the change into conformation II. Neutralizing three charges of the cluster (3\*) induced significant differences as seen in the <sup>15</sup>N-HSQC spectrum at neutral pH conditions (Fig. 5d and Supplementary Fig. S7). Triple-resonance NMR spectroscopy showed that the overall structure of L.h. 3\* (Supplementary Fig. S2, pdb 2N3E) resembled the five-helix bundle of NRN1<sub>E.a.</sub><sup>12,28</sup>. A detailed triple resonance assignment is given in Schaal *et al.*<sup>29</sup>. In summary it can be assumed that the charges of the acidic cluster are not required to adopt the overall structure of wtNRN1<sub>L.h.</sub>, but they significantly affect conformation I. The rearrangement into conformation II seems to be initiated by neutralization of the clustered acidic residue side chains. Protonation of additional residues might further have an impact on the conformation, as suggested by a slightly changed <sup>15</sup>N-HSQC fingerprint of 3\* upon acidification (Fig. 5d).

The structural conversion of NRN1<sub>L.h.</sub> was independent of dimerization, and both were controlled by distinct amino acids of the domain. Protonation of E114 apparently triggers the antiparallel dimerization by folding into a slightly altered conformation I\*, while the conversion to conformation II is mediated by residues of the acidic cluster (Fig. 6).



## Discussion

The dimerization of the amino-terminal spider silk domain NRN1<sub>Lh</sub> is triggered by conserved acidic residues in a pH- and salt-dependent manner (Fig. 6). During spidroin storage in the spider's sac, a neutral pH and the presence of chaotropic salts like NaCl stabilize NRN1<sub>Lh</sub> in the monomeric conformation I. The decrease of sodium chloride concentration along the silk gland supports the formation of electrostatic interactions between the acidic and basic poles of each monomer, prearranging the antiparallel dimer. Coincident with a dimerization mechanism suggested for NRN1<sub>Ea</sub><sup>10</sup>, the binding is predominantly stabilized by electrostatic interactions involving residues D39 and E81. Acidification along the spinning duct triggers neutralization of residue E114 and, thereby, slightly changes into conformation I\* to strengthen the dimer formation. However, protonation of E76 does not trigger dimerization of NRN1<sub>Lh</sub> as it was suggested for NRN1<sub>Ea</sub><sup>10</sup>.

Successive protonation of the structurally neighbouring acidic residues D39, E76 and E81 reduces the charge repulsion between helix 2 and 3 and rearranges the five-helix bundle into the tight conformation II. This mechanism slightly differs from that of NRN1<sub>Ea</sub> in which apparently only E81 plays a key role<sup>10</sup>. The acidic cluster ensures the correct order of firstly prearranging the dimer by electrostatic interactions and secondly changing the conformation of the dimerized protein.

So far, it is unclear if D39 and E81 get protonated and, thereby, reduce the binding strength of the dimer by breaking intermolecular electrostatic interactions, even so the monomer-dimer distribution remains unaffected at low pH values (Fig. 4a). In accordance with da Silva *et al.* D39 likely has a strong tendency to lose its proton upon dimerization of NRN1<sub>Ea</sub>, but neutralization did not affect dimer formation<sup>44</sup>. In spiders, the pH values were only determined for one species in the first half of the spinning duct due to its extremely small inner diameter making it difficult to speculate about physiologically relevant protonation of residues<sup>11</sup>. In case of further acidification in the second half of the duct its function remains unresolved since a stable NRN dimer is obviously formed at pH 6.0, which is already present in the middle part of the duct.

The sequence of the amino-terminal domain, especially the acidic amino acids at positions 39, 76, 81 and 114, is highly conserved between different spiders<sup>24</sup>. Therefore, it is not surprising that the dimerization mechanism is similar between species. However, the key residues that control the step-by-step dimerization slightly diverge between homologue amino-terminal domains, comparable to discrepancies found between major and minor ampullate spidroins in one species<sup>24</sup>. In case of artificial fiber spinning it will be necessary to identify the molecule-specific tightly controlled process to obtain fibers with a toughness identical to that of natural ones. Since the amino-terminal domain allows to increase the extensibility and strength of fibers made of recombinant spider silk proteins<sup>19</sup>, the stepwise dimerization of their amino-terminal domain is a necessary prerequisite for correct spidroin assembly.

## References

- Xu, M. & Lewis, R. V. Structure of a protein superfiber: spider dragline silk. *Proc. Natl. Acad. Sci. USA* **87**, 7120–7124 (1990).
- Hinman, M. B. & Lewis, R. V. Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber. *J. Biol. Chem.* **267**, 19320–19324 (1992).
- Ayoub, N. A., Garb, J. E., Tinghitella, R. M., Collin, M. A. & Hayashi, C. Y. Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. *Plos One* **2**, e514 (2007).
- Hijirida, D. H. *et al.* 13C NMR of *Nephila clavipes* major ampullate silk gland. *Biophys. J.* **71**, 3442–3447 (1996).
- Chen, X., Knight, D. P. & Vollrath, F. Rheological characterization of *Nephila* spidroin solution. *Biomacromolecules* **3**, 644–648 (2002).
- Gosline, J. M., Guerette, P. A., Ortlepp, C. S. & Savage, K. N. The mechanical design of spider silks: From fibroin sequence to mechanical function. *J. Exp. Biol.* **202**, 3295–3303 (1999).
- Knight, D. P. & Vollrath, F. Changes in element composition along the spinning duct in a *Nephila* spider. *Naturwissenschaften* **88**, 179–182 (2001).
- Dicko, C., Vollrath, F. & Kenney, J. M. Spider silk protein refolding is controlled by changing pH. *Biomacromolecules* **5**, 704–710 (2004).
- Foo, C. W. P. *et al.* Role of pH and charge on silk protein assembly in insects and spiders. *Appl. Phys. A: Mater* **82**, 223–233 (2006).
- Kronqvist, N. *et al.* Sequential pH-driven dimerization and stabilization of the N-terminal domain enables rapid spider silk formation. *Nat. Commun.* **5**, 3254 (2014).
- Andersson, M. *et al.* Carbonic Anhydrase Generates CO<sub>2</sub> and H<sup>+</sup> That Drive Spider Silk Formation Via Opposite Effects on the Terminal Domains. *Plos Biol.* **12**, e1001921 (2014).
- Askarieh, G. *et al.* Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature* **465**, 236–238 (2010).
- Hagn, F. *et al.* A highly conserved spider silk domain acts as a molecular switch that controls fibre assembly. *Nature* **465**, 239–242 (2010).
- Simmons, A. H., Michal, C. A. & Jelinski, L. W. Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk. *Science* **271**, 84–87 (1996).
- Huemmerich, D. *et al.* Primary structure elements of spider dragline silks and their contribution to protein solubility. *Biochemistry* **43**, 13604–13612 (2004).
- Exler, J. H., Hummerich, D. & Scheibel, T. The amphiphilic properties of spider silks are important for spinning. *Angew. Chem. Int. Ed. Engl.* **46**, 3559–3562 (2007).
- Heim, M., Keerl, D. & Scheibel, T. Spider silk: from soluble protein to extraordinary fiber. *Angew. Chem. Int. Ed. Engl.* **48**, 3584–3596 (2009).
- Eisoldt, L., Smith, A. & Scheibel, T. Decoding the secrets of spider silk. *Mater. Today* **14**, 80–86 (2011).
- Heidebrecht, A. *et al.* Biomimetic Fibers Made of Recombinant Spidroins with the Same Toughness as Natural Spider Silk. *Adv. Mater.* **27**, 2189–2194 (2015).
- Bini, E., Knight, D. P. & Kaplan, D. L. Mapping domain structures in silks from insects and spiders related to protein assembly. *J. Mol. Biol.* **335**, 27–40 (2004).
- Rising, A., Hjalms, G., Engstrom, W. & Johansson, J. N-terminal nonrepetitive domain common to dragline, flagelliform, and cylindriciform spider silk proteins. *Biomacromolecules* **7**, 3120–3124 (2006).
- Garb, J. E., Ayoub, N. A. & Hayashi, C. Y. Untangling spider silk evolution with spidroin terminal domains. *Bmc Evol. Biol.* **10**, 243 (2010).
- Chen, G. F. *et al.* Full-Length Minor Ampullate Spidroin Gene Sequence. *Plos One* **7**, e52293 (2012).

24. Otkovs, M. *et al.* Diversified Structural Basis of a Conserved Molecular Mechanism for pH-Dependent Dimerization in Spider Silk N-Terminal Domains. *ChemBiochem* **16**, 1720–1724 (2015).
25. Gaines, W. A., Sehorn, M. G. & Marcotte, W. R. Spidroin N-terminal domain promotes a pH-dependent association of silk proteins during self-assembly. *J. Biol. Chem.* **285**, 40745–40753 (2010).
26. Hagn, F., Thamm, C., Scheibel, T. & Kessler, H. pH-Dependent Dimerization and Salt-Dependent Stabilization of the N-terminal Domain of Spider Dragline Silk—Implications for Fiber Formation. *Angew. Chem. Int. Ed. Engl.* **50**, 310–313 (2011).
27. Landreh, M. *et al.* A pH-dependent dimer lock in spider silk protein. *J. Mol. Biol.* **404**, 328–336 (2010).
28. Jaudzems, K. *et al.* pH-dependent dimerization of spider silk N-terminal domain requires relocation of a wedged tryptophan side chain. *J. Mol. Biol.* **422**, 477–487 (2012).
29. Schaaf, D. *et al.* Resonance assignment of an engineered amino-terminal domain of a major ampullate spider silk with neutralized charge cluster. *Biomol. NMR Assign.* **10**, 199–202 (2016).
30. Studier, F. W. Protein production by auto-induction in high-density shaking cultures. *Protein Express. Purif.* **41**, 207–234 (2005).
31. Savitzky, A. & Golay, M. J. E. Smoothing + Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **36**, 1627–1639 (1964).
32. Mori, S., Abeygunawardana, C., Johnson, M. O. & Vanzijl, P. C. M. Improved Sensitivity of HsQC Spectra of Exchanging Protons at Short Interscan Delays Using a New Fast HsQC (FhsQC) Detection Scheme That Avoids Water Saturation. *J. Magn. Reson., Ser B* **108**, 94–98 (1995).
33. Kay, L. E., Torchia, D. A. & Bax, A. Backbone Dynamics of Proteins as Studied by N-15 Inverse Detected Heteronuclear NMR Spectroscopy - Application to Staphylococcal Nuclease. *Biochemistry* **28**, 8972–8979 (1989).
34. Rossi, P. *et al.* A microscale protein NMR sample screening pipeline. *J. Biomol. NMR* **46**, 11–22 (2010).
35. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic. Acids. Res.* **28**, 235–242 (2000).
36. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73 (2003).
37. Schwieters, C. D., Kuszewski, J. J. & Clore, G. M. Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Mag. Res. Sp.* **48**, 47–62 (2006).
38. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
39. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* **8**, 477–486 (1996).
40. Vranken, W. F. *et al.* The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **59**, 687–696 (2005).
41. DeLano, W. L. *The PyMOL Molecular Graphics System, version 1, 3r1* (2010).
42. Schwarze, S., Zwettler, F. U., Johnson, C. M. & Neuweiler, H. The N-terminal domains of spider silk proteins assemble ultrafast and protected from charge screening. *Nat. Commun.* **4**, 2815 (2013).
43. Wallace, J. A. & Shen, J. K. Unraveling a trap-and-trigger mechanism in the pH-sensitive self-assembly of spider silk proteins. *J. Phys. Chem.* **3**, 658–662 (2012).
44. da Silva, F. L. B., Pasquali, S., Derreumaux, P. & Dia, L. G. Electrostatic analysis of the mutational and pH effects of the N-terminal domain self-association of the major ampullate spidroin. *Soft Matter* **12**, 5600–5612 (2016).

## Acknowledgements

We are grateful to Dr. Martin Humenik for experimental help and Britta Zimmermann for performing the fermentation of *E. coli* to produce <sup>15</sup>N-labelled proteins. Funding was obtained from the DFG (SFB 840 TP A8).

## Author Contributions

J.B., D.S., L.E., S.S. and T.S. conceived the study design. J.B. and D.S. conducted the experiments. K.S., S.S. and T.S. supervised the research. J.B. wrote the manuscript with input from all authors. D.S., S.S. and T.S. revised the article critically.

## Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Bauer, J. *et al.* Acidic Residues Control the Dimerization of the N-terminal Domain of Black Widow Spiders' Major Ampullate Spidroin 1. *Sci. Rep.* **6**, 34442; doi: 10.1038/srep34442 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016

# Supplementary Information

## Acidic Residues Control the Dimerization of the N-terminal Domain of Black Widow Spiders' Major Ampullate Spidroin 1

Joschka Bauer<sup>1</sup>, Daniel Schaal<sup>2,3</sup>, Lukas Eisoldt<sup>1</sup>, Kristian Schweimer<sup>2,3</sup>, Stephan Schwarzinger<sup>2,3</sup>, and Thomas Scheibel<sup>1,3,4,5,6,\*</sup>

<sup>1</sup> Lehrstuhl Biomaterialien, Universität Bayreuth, Fakultät für Ingenieurwissenschaften, Bayreuth, Germany

<sup>2</sup> Lehrstuhl Biopolymere, Universität Bayreuth, Bayrisches Geoinstitut, Bayreuth, Germany

<sup>3</sup> Forschungszentrum für Bio-Makromoleküle (BIOMac), Bayrisches Geoinstitut, Universität Bayreuth, Bayreuth, Germany

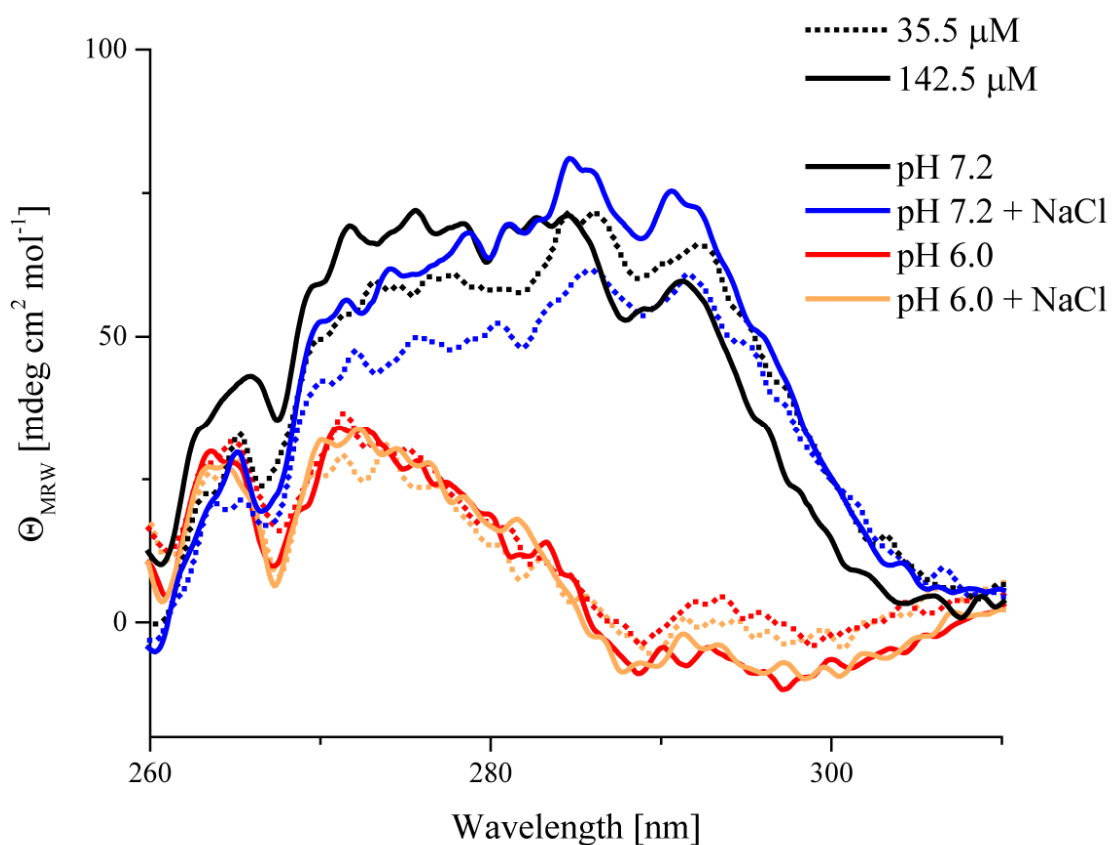
<sup>4</sup> Bayreuther Zentrum für Kolloide und Grenzflächen (BZKG), Naturwissenschaften I, Universität Bayreuth, Bayreuth, Germany

<sup>5</sup> Bayreuther Materialzentrum (BayMat), Fakultät für Ingenieurwissenschaften, Universität Bayreuth, Bayreuth, Germany

<sup>6</sup> Bayreuther Zentrum für Molekulare Biowissenschaften (BZMB), Naturwissenschaften I, Universität Bayreuth, Bayreuth, Germany

\* Corresponding author

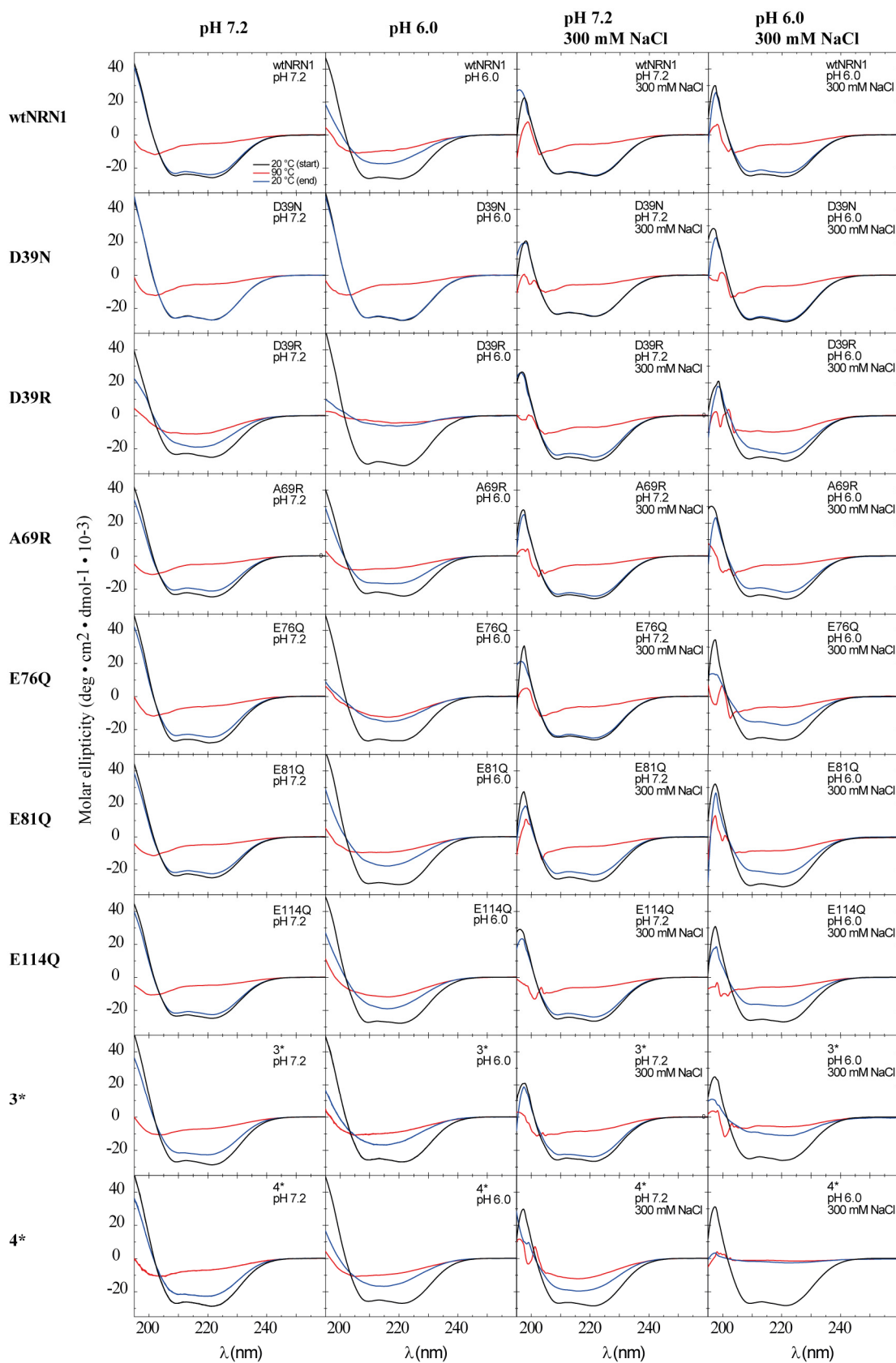
E-mail: [thomas.scheibel@bm.uni-bayreuth.de](mailto:thomas.scheibel@bm.uni-bayreuth.de)



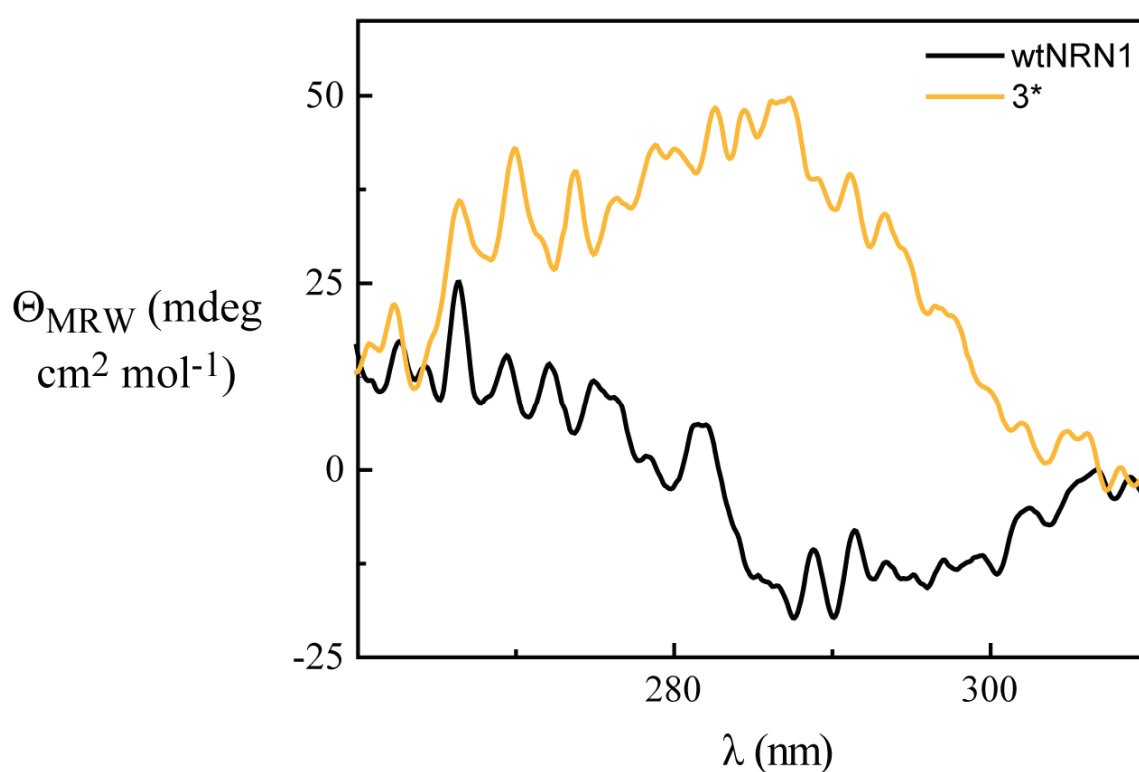
**Figure S1. Effect of wtNRN1<sub>L.h.</sub> concentration on near-UV CD spectra.** Near-UV CD spectra were taken at a wtNRN1<sub>L.h.</sub> concentration of 142.5  $\mu$ M using a path length of 0.5 cm (solid line) or at 35.5  $\mu$ M using a path length of 1.0 cm (dotted). Spectra were measured at pH 7.2 (black), pH 7.2 in the presence of 300 mM NaCl (blue), pH 6.0 (red) and pH 6.0 in the presence of 300 mM NaCl (orange).



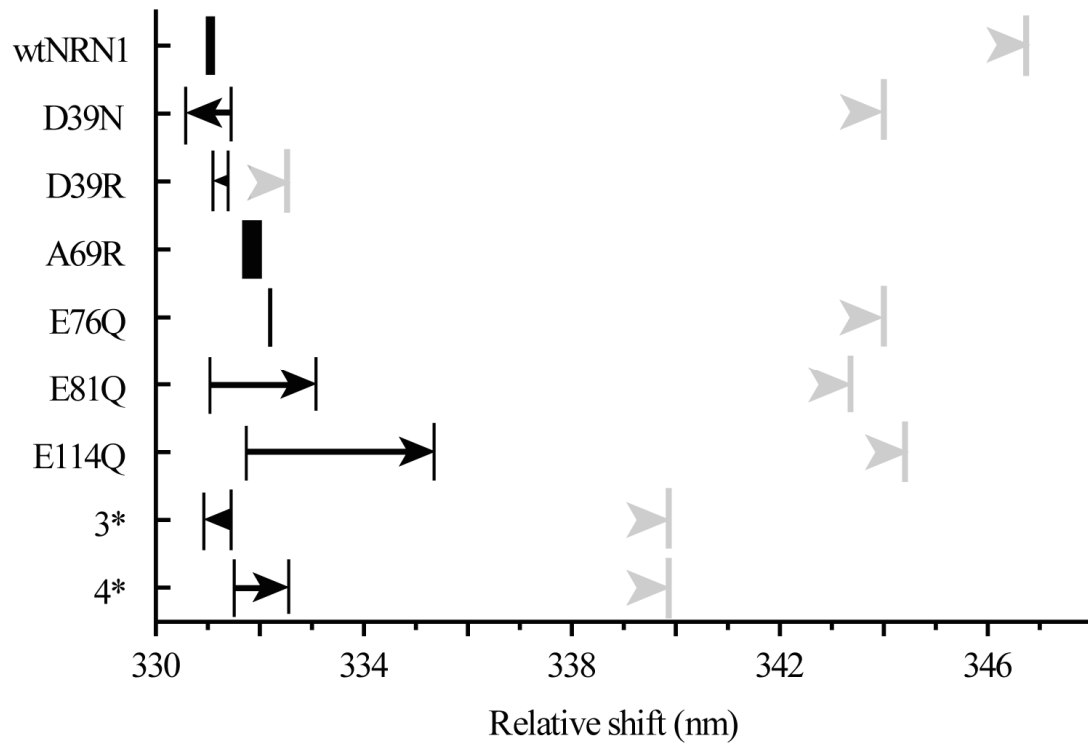
**Figure S2. Backbone overlay of the twenty lowest energy structures of NRN1<sub>L.h.</sub> 3\*.** The root mean square deviation to the mean structure amounts for 0.53 Å (backbone heavy atoms).



**Figure S3. Far-UV CD spectra of NRN1<sub>L.h.</sub> variants.** The residual molar ellipticity was initially determined at 20 °C (black line). Afterwards, samples were heated to 90 °C (red) and again cooled to 20 °C (blue).

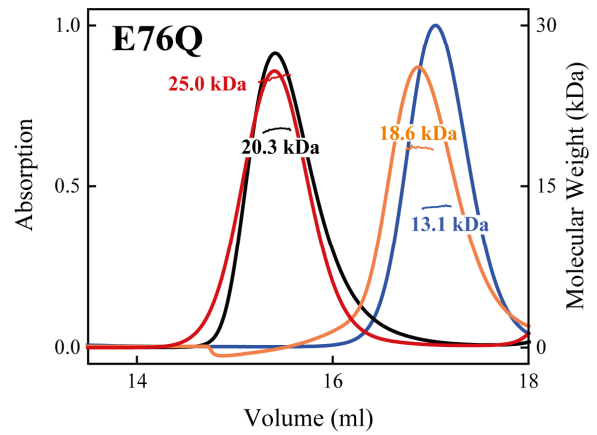
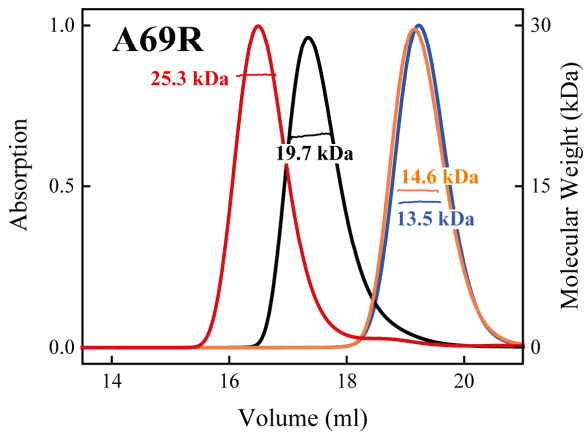
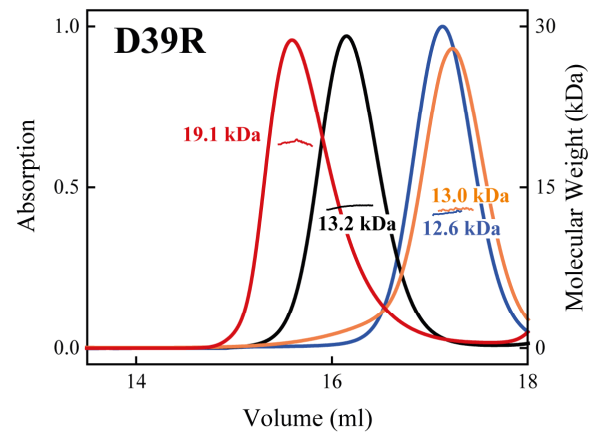
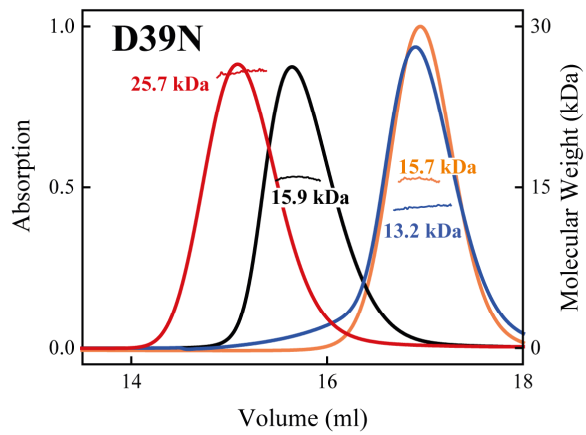
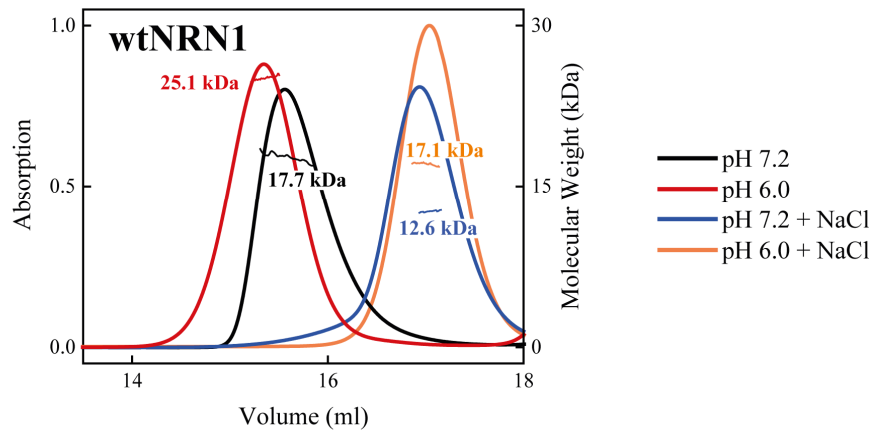


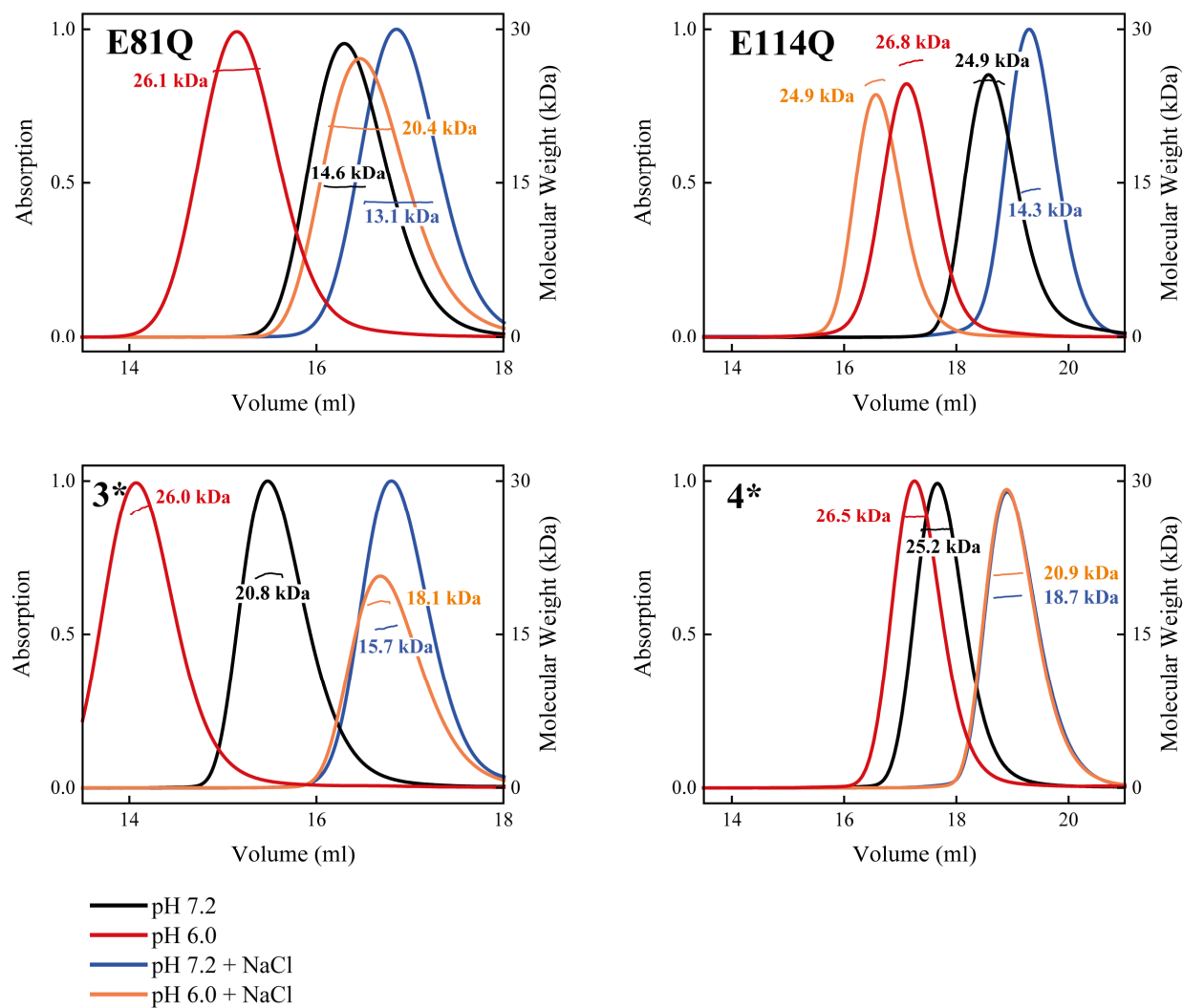
**Figure S4. NaCl restricted the structural conversion of variant 3\* compared to that of wtNRN1<sub>L.h.</sub>** The difference of the near-UV CD spectra between pH 7.2 and pH 6.0 (see also Figs. 1b and d) in the absence of salt are subtracted by the differential spectra in the presence of 300 mM NaCl (see also Figs. 1c and e). While the presence of NaCl only marginally affected the conformational change of wtNRN1<sub>L.h.</sub>, the structural flexibility of 3\* was significantly suppressed.



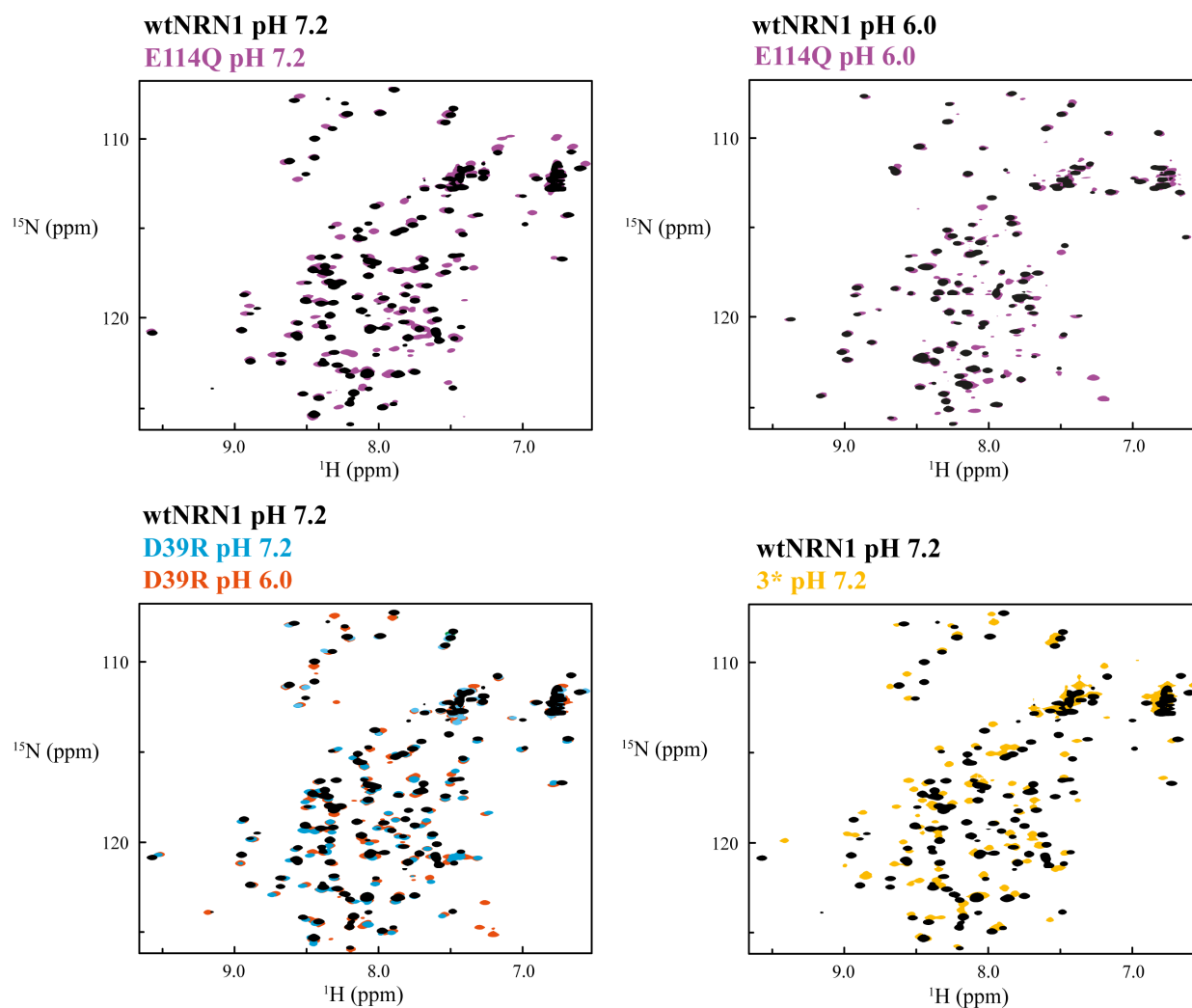
**Figure S5. Shift of the fluorescence maximum upon lowering the pH from 7.2 to 6.0 in the presence of sodium chloride.** 300 mM NaCl (black arrow) reduces the fluorescence redshift compared to the fluorescence maximum in the absence of salt at pH 7.2 (grey arrow, see also Fig. 2b).



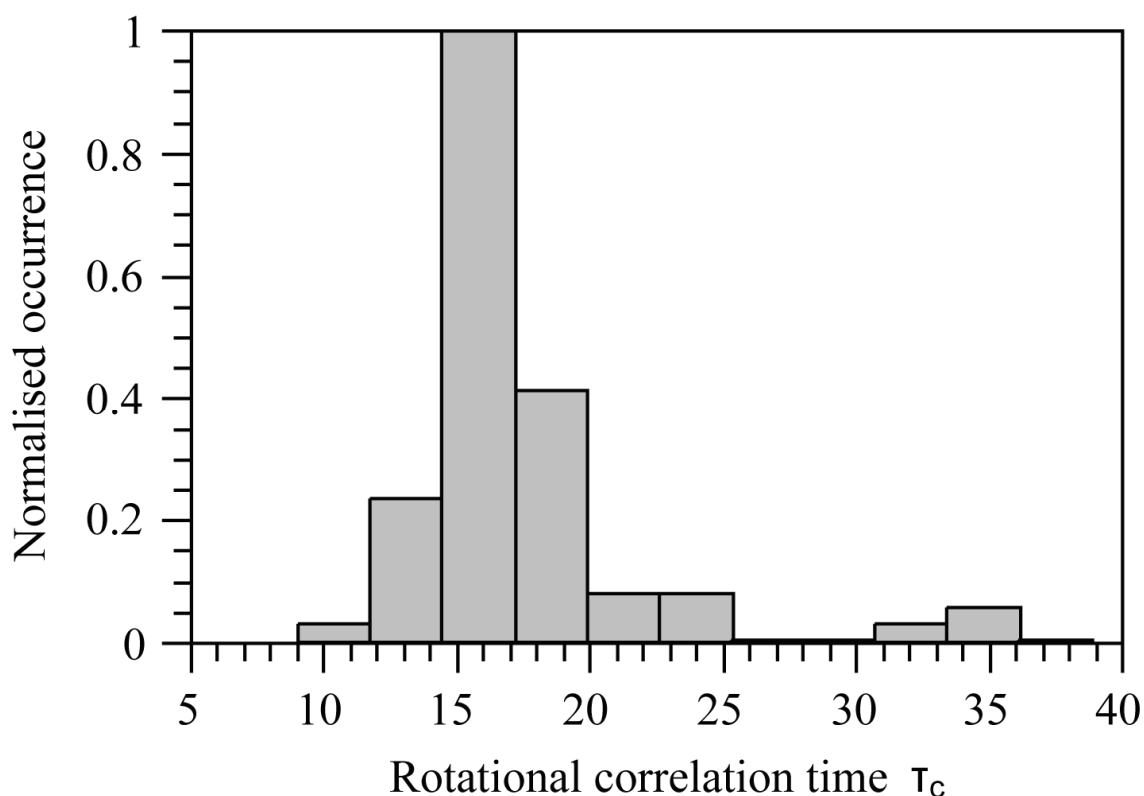




**Figure S6. Chromatograms of wildtype as well as variants of NRN1<sub>L.h.</sub>.** The chromatograms show the UV signal, and the molar mass of the proteins is indicated. MALS was measured at pH 7.2 (black) as well as pH 6.0 (red) in the absence and presence of 300 mM NaCl (pH 7.2 + NaCl, blue; pH 6.0 + NaCl, orange).



**Figure S7: Overlay of HSQC spectra of wtNRN1<sub>L.h.</sub> and variants thereof.** For a better comparison of  $^{15}\text{N}$ -HSQC fingerprints, the spectrum of wtNRN1<sub>L.h.</sub> (black) was overlaid with spectra of E114Q (purple), D39R at pH 7.2 (blue) and pH 6.0 (red) as well as 3\* at pH 7.2 (orange).



**Figure S8. Rotational correlation time  $\tau_c$  of E114Q at pH 7.2.** The rotational correlation

time  $\tau_c$  was estimated using the equation  $\tau_c = \sqrt{\frac{6T_1}{T_2} - 7} \frac{1}{4\pi\nu_N}$  with  $T_1$  and  $T_2$  as the

longitudinal and transversal relaxation times (delays used in the acquisition of the relaxation

data are listed in Supplementary Table S3) and  $\nu_N$  as the  $^{15}\text{N}$  nuclear frequency, respectively<sup>1</sup>.

The distribution of the rotational correlation times of the individual sequence positions

showed an unimodal pattern with the highest frequency at around 16-17 ns, which is typical

for molecular weights of around 25-30 kDa<sup>2</sup>, confirming that the variant E114Q dimerizes at a physiological pH.

**Table S1. Nuclear overhauser effect (NOE) restraints and structural statistics for the best 20 structures of NRN1<sub>L.h.</sub> variant 3\*.**

<b>Experimental derived restraints</b>		
distance restraints	NOE	1253
	intraresidual	9
	sequential	466
	medium range	413
	long range	365
dihedral restraints		224
<b>restraint violation</b>		
average distance restraint violation (Å)		0.008118 (+/- 0.000900)
distance restraint violation > 0.1 Å		0.95 (+/- 1.20)
distance restraint violation > 0.3 Å		0.00 (+/- 0.00)
average dihedral restraint violation (°)		0.3799 (+/- 0.0332)
dihedral restraint violation > 1°		8.45 (+/- 1.63)
dihedral restraint violation > 3°		0.05 (+/- 0.22)
dihedral restraint violation > 5°		0.00 (+/- 0.00)
<b>deviation from ideal geometry</b>		
bond length (Å)		0.000863 (+/- 0.000058)
bond angle (°)		0.1715 (+/- 0.0084)
<b>coordinate precision<sup>a,b</sup></b>		
backbone heavy atoms (Å)		0.53
all heavy atoms (Å)		0.89
<b>Ramachandran plot statistics<sup>c</sup> (%)</b>		
	regions	
	most favoured	93.3
	additionally allowed	6.2
	generously allowed	0.4
	disallowed	0.0

<sup>a</sup> The precision of the coordinates is defined as the average atomic root mean square difference between the accepted simulated annealing structures and the corresponding mean structure calculated for the given sequence regions.

<sup>b</sup> calculated for residues Ala5-Ala127

<sup>c</sup> Ramachandran plot statistics are determined by PROCHECK (Laskowski, et al., 1993) and noted by most favored/additionally allowed/generously allowed/disallowed.

**Table S2. Stability of NRN1<sub>L.h.</sub> variants.** Protein denaturation was analysed as a function of urea concentration or temperature using CD ellipticity at 222 nm. The urea concentration and melting temperature ( $T_m$ ) refer to the condition at which half of the proteins are denatured. The min/max values at distinct solvent conditions are highlighted in blue/red, and equally stable variants are marked by stars (\*).

	Urea (M)		T <sub>m</sub> (°C)			
	pH 7.2	pH 6.0	pH 7.2	pH 7.2 + 300 mM NaCl	pH 6.0	pH 6.0 + 300 mM NaCl
<b>wtNRN1</b>	1.8	1.9	47.4	59.8	65.5*	62.1
<b>D39N</b>	2.9	3.0	53.3	63.3	54.2	63.5
<b>D39R</b>	4.2*	4.4*	61.2	71.3	65.6*	69.8
<b>A69R</b>	1.7	2.3	42.4	54.9	45.8	57.2
<b>E76Q</b>	2.4	2.4	55.0	62.9	65.6*	64.8
<b>E81Q</b>	2.4	2.4	49.3	59.8	62.6	61.5
<b>E114Q</b>	2.4	2.7	63.0	60.3	64.9	63.0
<b>3*</b>	3.5	3.4	60.2	65.7	61.0	67.3
<b>4*</b>	3.7	3.8	64.7	65.1	62.7	68.4

**Table S3.  $T_1$  and  $T_2$  relaxation delays used in the acquisition of  $^{15}\text{N}$ -HSQC spectra of variant E114Q.**

$T_2$ relaxation delays (ms)	$T_1$ relaxation delays (ms)
8.48	10.642
8.48	10.642
16.96	265.57
25.44	265.57
25.44	637.34
33.92	637.34
42.4	1,062.22
42.4	1,062.22
50.88	1,380.88
59.36	1,380.88
59.36	1,911.98
67.84	1,911.98
84.8	2,655.52
84.8	2,655.52
93.28	
93.28	

## References

- 1 Kay, L. E., Torchia, D. A. & Bax, A. Backbone Dynamics of Proteins as Studied by N-15 Inverse Detected Heteronuclear Nmr-Spectroscopy - Application to Staphylococcal Nuclease. *Biochemistry* **28**, 8972-8979, (1989).
- 2 Rossi, P. *et al.* A microscale protein NMR sample screening pipeline. *J. Biomol. NMR* **46**, 11-22, (2010).

Teilarbeit II erschien 2016 in *Biomolecular NMR Assignments*.

Schaal, D.; Bauer, J.; Schweimer, K.; Scheibel, T. & Schwarzing, S. (2016) Resonance assignment of an engineered amino-terminal domain of a major ampullate spider silk with neutralized charge cluster. *Biomolecular NMR Assignments* 1: 199.

Im Folgenden wird die Publikation mit Genehmigung des Verlages Springer Science und Business Media reproduziert.



# *Resonance assignment of an engineered amino-terminal domain of a major ampullate spider silk with neutralized charge cluster*

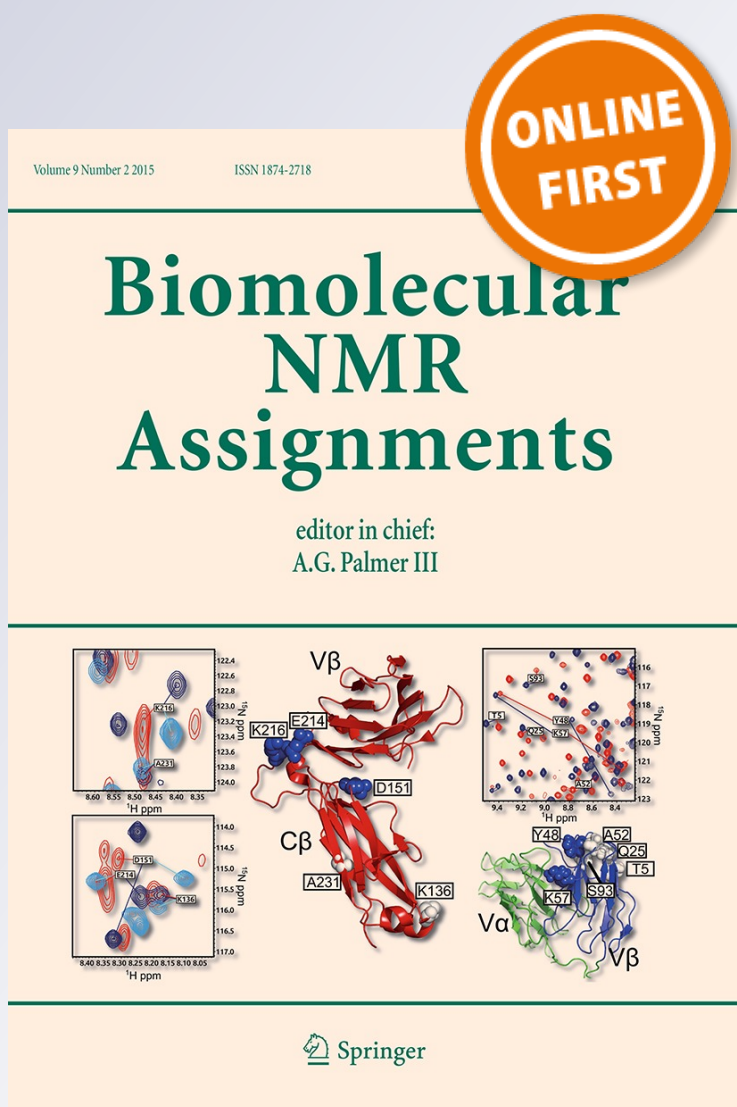
**Daniel Schaal, Joschka Bauer, Kristian Schweimer, Thomas Scheibel, Paul Rösch & Stephan Schwarzhinger**

**Biomolecular NMR Assignments**

ISSN 1874-2718

Biomol NMR Assign

DOI 10.1007/s12104-016-9666-y



**Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at [link.springer.com](http://link.springer.com)".**

# Resonance assignment of an engineered amino-terminal domain of a major ampullate spider silk with neutralized charge cluster

Daniel Schaal<sup>2,3</sup> · Joschka Bauer<sup>1</sup> · Kristian Schweimer<sup>2,3</sup> · Thomas Scheibel<sup>1,3,4,5,6</sup> · Paul Rösch<sup>2,3</sup> · Stephan Schwarzingen<sup>2,3</sup>

Received: 25 September 2015 / Accepted: 5 February 2016  
© Springer Science+Business Media Dordrecht 2016

**Abstract** Spider dragline fibers are predominantly made out of the major ampullate spidroins (MaSp) 1 and 2. The assembly of dissolved spidroin into a stable fiber is highly controlled for example by dimerization of its amino-terminal domain (NRN) upon acidification, as well as removal of sodium chloride along the spinning duct. Clustered residues D39, E76 and E81 are the most highly conserved residues of the five-helix bundle, and they are hypothesized to be key residues for switching between a monomeric and a dimeric conformation. Simultaneous replacement of these residues by their non-titratable analogues results in variant D39N/E76Q/E81Q, which is supposed to fold into an intermediate conformation between that of the monomeric and the dimeric state at neutral pH. Here we report the resonance assignment of *Latrodectus hesperus* NRN

variant D39N/E76Q/E81Q at pH 7.2 obtained by high-resolution triple resonance NMR spectroscopy.

**Keywords** Spider silk · Amino-terminal domain · Major ampullate spidroin 1 · Dimerization · *Latrodectus hesperus* · Acidic charge cluster

## Biological context

The terminal domains of major ampullate spidroins control the assembly of dragline spider silks intermolecular protein contacts (Hinman and Lewis 1992; Xu and Lewis 1990). Upon lowering the pH and decreasing the sodium chloride concentration of the protein solution along the spinning duct, the amino-terminal domain forms a homodimer and, thereby, initiates spidroin assembly (Gaines et al. 2010; Hagn et al. 2010; Landreh et al. 2010). The sequence of NRN is highly conserved between different silk types and spider species, underlining the crucial role of NRN for controlling spidroin assembly (Bini et al. 2004; Chen et al. 2012; Garb et al. 2010; Rising et al. 2006). X-ray crystal and NMR solution structures of *Euprosthenops australis* NRN revealed separation of acidic and basic amino acids, prearranging the dimer in antiparallel (Askarieh et al. 2010). The negative charge cluster is composed of three highly conserved residues—D39, E76 and E81—that are hypothesized to control the pH-dependent dimerization and the simultaneous conversion into a tighter conformation, as evidenced by NMR and crystals structures. Neutralization of the acidic charge cluster suppresses the electrostatic repulsion between helices 2 and 3, which leads to their rearrangement and subsequent flattening of the dimerization interface (Bini et al. 2004; Kronqvist et al. 2014). To study the role of the acidic cluster during the

✉ Daniel Schaal  
daniel.schaal@uni-bayreuth.de;  
<http://www.biopolymere.uni-bayreuth.de>

<sup>1</sup> Lehrstuhl Biomaterialien, Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

<sup>2</sup> Lehrstuhl Biopolymere, Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

<sup>3</sup> Forschungszentrum für Bio-Makromoleküle (BIOmac), Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

<sup>4</sup> Bayreuther Zentrum für Kolloide und Grenzflächen (BZKG), Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

<sup>5</sup> Bayreuther Materialzentrum (BayMat), Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

<sup>6</sup> Bayreuther Zentrum für Molekulare Biowissenschaften (BZMB), Universität Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

conformational change, the clustered aspartic and glutamic acid residues were simultaneously substituted by their non-titratable analogues asparagine and glutamine, resulting in the NRN variant D39N/E76Q/E81Q. This variant mimics a protonated state of NRN at low pH, which reflects conditions as found close to the end of the spider's spinneret. The majority of published research was done on *E. australis* MaSp1. For NRN from the black widow spider *Latrodectus hesperus*, solely monomeric wild type resonance assignments, but no structural coordinates or NMR distance restraints were published [Hagn et al. 2011, BioMagResBank (BMRB) accession code 17131]. Here we collected three-dimensional NMR data of D39N/E76Q/E81Q and assigned backbone as well as sidechain resonances.

## Methods and experiments

### Protein expression and purification

The variant *L.h.* MaSp1 NRN D39N/E76Q/E81Q was obtained by cloning MaSp1-NRN cDNA—mutated by using QuikChange® Site-Directed Mutagenesis Kit (Agilent, Santa Clara, CA, US)—together with a His<sub>6</sub>-SUMO-tag in vector pET28a (Novagen, Merck, Darmstadt, Germany). Genes were transformed to and expressed in *Escherichia coli* BL21 (DE3). Before IPTG induction the cells were grown to an OD<sub>600</sub> = 0.7 in M9 minimal medium at 37 °C containing kanamycin, <sup>15</sup>N-ammonium sulphate and <sup>13</sup>C-glucose as exclusive nitrogen and carbon source, respectively. After 5 h of protein production, the cells were harvested by centrifugation for 12 min at 12,100g and 4 °C. Cells were opened using a Microfluidizer M-110S (Microfluidics, Westwood, MA, US) at 6.5 bar twice. Proteins were purified by Ni-NTA chromatography (HisTrap FF, GE Healthcare, Little Chalfont, UK) and size exclusion chromatography (HiLoad™ 26/60 Superdex™ 75 pg, GE Healthcare). After SUMO protease cleavage another Ni-NTA chromatography was performed to separate the tag from the protein. The protein was freeze-dried and stored at −20 °C.

### NMR experiments

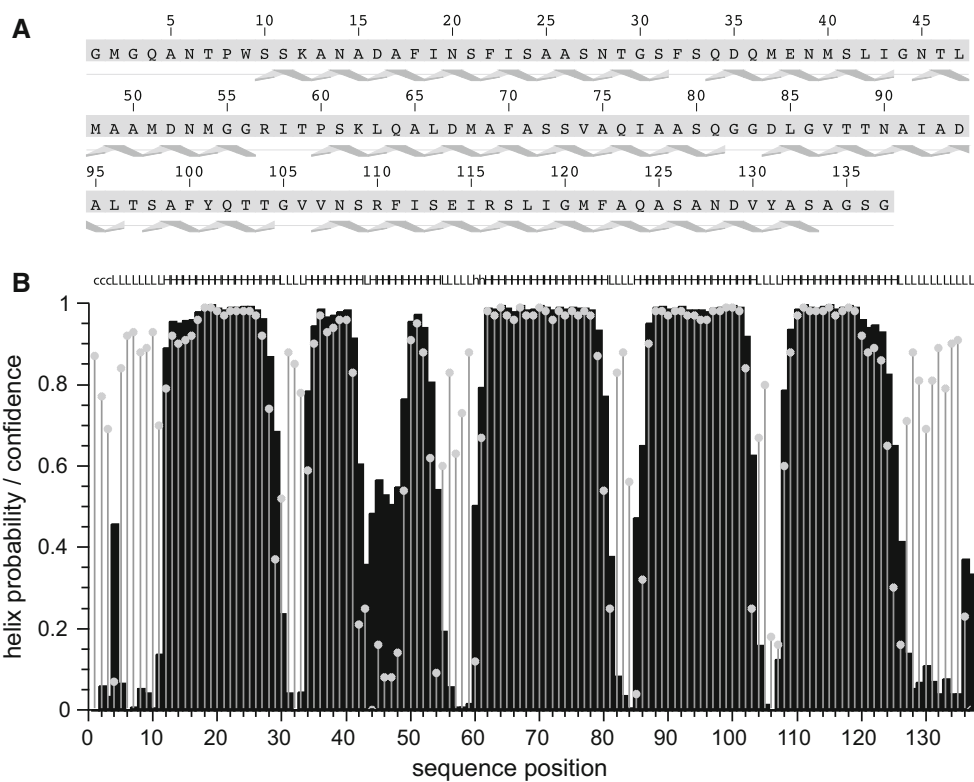
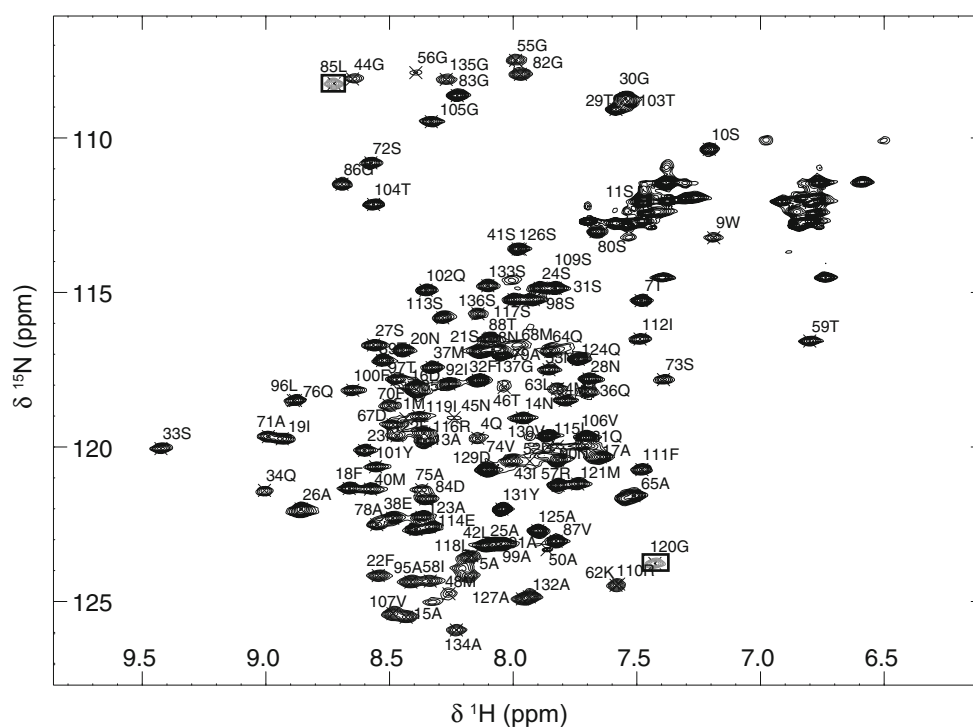
The NMR samples were prepared by dissolving the freeze-dried protein in 22 mM sodium phosphate buffer at a concentration of 0.6 mM, the addition of 10 % (v/v) D<sub>2</sub>O and final pH adjustment to 7.2. All NMR data were recorded on a Bruker Avance 700 and Avance II+ 600 MHz NMR spectrometer equipped with a 5 mm TCI cryo and TXI probe with Z-axis gradients, respectively. For sequential backbone assignment standard HNCA, HNCACB, CBCA(CO)NH and

<sup>15</sup>N-resolved NOESY-HSQC spectra were recorded (Sattler et al. 1999). For side chain assignment <sup>13</sup>C-resolved aliphatic and aromatic NOESY-HSQC, as well as (H)CCH-TOCSY and NOESY spectra were acquired (Marion et al. 1989; Zuiderweg and Fesik 1989). Additionally two-dimensional <sup>15</sup>N- and <sup>13</sup>C-HSQC experiments were recorded regularly to check for protein stability (Mori et al. 1995). Spectral analysis, resonance assignment and imaging was done with the CCPNMR software package (Vranken et al. 2005), NMRViewJ (One Moon Scientific, Westfield, NJ, US), qtiplot (IONDEV SRL, Bucurest, Romania) and Adobe Illustrator CS3 (Adobe Systems, San Jose, CA, US).

### Resonance assignment, secondary structure prediction and data deposition

The <sup>15</sup>N-HSQC spectrum of D39N/E76Q/E81Q showed well-dispersed signals as typical for a well-folded protein (Fig. 1). TALOS-N was used to perform a secondary structure estimation that indicated high helix probability for regions 12–29, 34–42, 44–54, 61–80, 85–103 and 108–125 (Shen and Bax 2013); Fig. 2). Sequential backbone assignment using standard experiments could be achieved for the majority of the chain. In sequence regions in helices 2 and 3 (39–78), where signal intensities were very weak, the assignments turned out to be more challenging. Additionally Met2, Pro8, Ile43, Pro60 and Ser61 could not be assigned. The dimerization interface could be anticipated by highlighting of region 39–78 on the homologue dimeric NRN from *E. australis* (PDB 2LTH, (Kronqvist et al. 2014). The modelled dimerization interface coincides with a sequence region of D39N/E76Q/E81Q that showed weak signal intensities in the NMR spectra, probably resulting from an intermediate chemical exchange of partial dimerization. C-terminal amino acids 134–137 neither showed inter-residual NOE nor sidechain proton signals. Through examination of all available types of spectra a total of 90.3 % of the backbone (HN, CA, N, HA) and 71.5 % of the total sidechain, 83 % of aromatic and 85.6 % of aliphatic side chain atoms could be assigned. <sup>13</sup>C<sup>α</sup> chemical shifts were compared to deposited wild type resonances (BioMagResBank accession code 17131) and summarized in Fig. 3. For most sequence positions of the protein <sup>13</sup>C<sup>α</sup> resonances of both variants were in good agreement. However some isolated residues showed unusual big shift differences. After careful re-evaluation of the assigned and deposited resonances with the recorded spectra and with the BMRB chemical shift statistics, some unusual chemical shifts could be observed for BMRB entry 17131. However, given that there is no access to the wild type spectra, a direct comparison of the wild type and variant three-dimensional data is currently not possible.

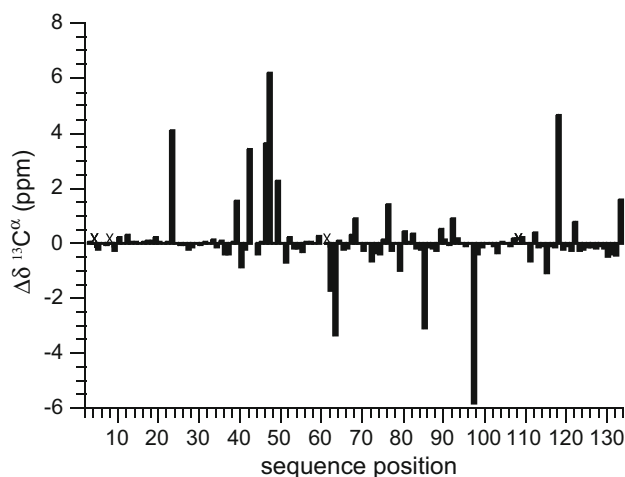
**Fig. 1**  $^{15}\text{N}$ -HSQC spectrum of NRN variant D39N/E76Q/E81Q in 20 mM sodium phosphate buffer with 10 % (v/v)  $\text{D}_2\text{O}$  at pH 7.2 and 298 K. Aliased signals are boxed



**Fig. 2** **a** Amino acid sequence of NRN variant D39N/E76Q/E81Q with predicted secondary structure from a chemical shift index using assigned HA, HN, CA resonances. **b** TALOS-N predicted helix probabilities (black bars) and secondary structure (top bar c = coil,

L = loop, H = helix), estimated from assigned chemical shifts with associated confidence levels as grey vertical lines (Shen and Bax 2013)





**Fig. 3** Chemical shift differences  $\Delta\delta^{13}\text{C}\alpha$  of deposited NRN wild type and variant D39N/E76Q/E81Q as a function of the sequence position. Unusual chemical shift differences are observed for individual residues and are addressed in the main text. Unassigned residues are marked with X

The resonances and assignments of NRN D39N/E76Q/E81Q have been deposited in the BioMagResBank (<http://www.bmrb.wisc.edu>) under the accession code 25643.

**Acknowledgments** Funding was obtained from the Deutsche Forschungsgemeinschaft DFG (SFB 840).

**Compliance with ethical standards**

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical standards** The authors declare that the experiments in this study comply with the current laws of the country in which they were performed.

## References

- Askarieh G et al (2010) Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature* 465:236–238
- Bini E, Knight DP, Kaplan DL (2004) Mapping domain structures in silks from insects and spiders related to protein assembly. *J Mol Biol* 335:27–40
- Chen GF et al (2012) Full-length minor ampullate spidroin gene sequence. *PLoS One* 7(12):e52293

- Gaines WA, Sehorn MG, Marcotte WR (2010) Spidroin N-terminal domain promotes a pH-dependent association of silk proteins during self-assembly. *J Biol Chem* 285:40745–40753
- Garb JE, Ayoub NA, Hayashi CY (2010) Untangling spider silk evolution with spidroin terminal domains. *BMC Evol Biol* 10:243
- Hagn F, Eisoldt L, Hardy JG, Vendrely C, Coles M, Scheibel T, Kessler H (2010) A highly conserved spider silk domain acts as a molecular switch that controls fibre assembly. *Nature* 465:239–242
- Hagn F, Thamm C, Scheibel T, Kessler H (2011) pH-dependent dimerization and salt-dependent stabilization of the N-terminal domain of spider dragline silk—implications for fiber formation. *Angew Chem Int Ed* 50:310–313
- Hinman MB, Lewis RV (1992) Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber. *J Biol Chem* 267:19320–19324
- Kronqvist N et al (2014) Sequential pH-driven dimerization and stabilization of the N-terminal domain enables rapid spider silk formation. *Nat Commun* 5:3254
- Landreh M et al (2010) A pH-dependent dimer lock in spider silk protein. *J Mol Biol* 404:328–336
- Marion D, Driscoll PC, Kay LE, Wingfield PT, Bax A, Gronenborn AM, Clore GM (1989) Overcoming the overlap problem in the assignment of  $^1\text{H}$  NMR spectra of larger proteins by use of three-dimensional heteronuclear  $^1\text{H}$ - $^{15}\text{N}$  Hartmann-Hahn-multiple quantum coherence and nuclear Overhauser-multiple quantum coherence spectroscopy: application to interleukin 1 beta. *Biochemistry* 28:6150–6156
- Mori S, Abeygunawardana C, Johnson MO, Vanzijl PCM (1995) Improved sensitivity of Hsqr spectra of exchanging protons at short interscan delays using a new fast Hsqr (Fhsqr) detection scheme that avoids water saturation. *J Magn Reson Ser B* 108:94–98
- Rising A, Hjalmar G, Engstrom W, Johansson J (2006) N-terminal nonrepetitive domain common to dragline, flagelliform, and cylindrical spider silk proteins. *Biomacromolecules* 7:3120–3124
- Sattler M, Schleucher J, Griesinger C (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog NMR Spectrosc* 34:93–158
- Shen Y, Bax A (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* 56:227–241
- Vranken WF et al (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* 59:687–696
- Xu M, Lewis RV (1990) Structure of a protein superfiber: spider dragline silk. *Proc Natl Acad Sci* 87:7120–7124
- Zuiderweg ER, Fesik SW (1989) Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a. *Biochemistry* 28:2387–2391

Teilarbeit III wurde 2017 bei *Biomacromolecules* publiziert.

Bauer, J. & Scheibel, T. (2017) Dimerization of the Conserved N-Terminal Domain of a Spider Silk Protein Controls the Self-Assembly of the Repetitive Core Domain. *Biomacromolecules* 18 (8): 2521.

Im Folgenden wird die Publikation mit Genehmigung mit Copyright © 2017 American Chemical Society reproduziert.

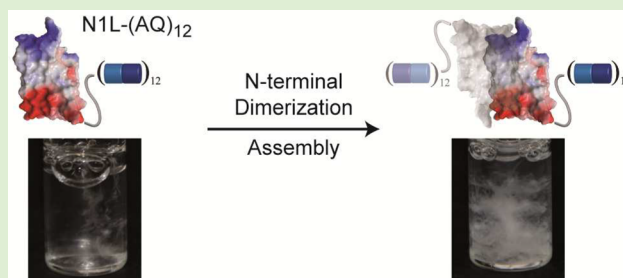
# Dimerization of the Conserved N-Terminal Domain of a Spider Silk Protein Controls the Self-Assembly of the Repetitive Core Domain

Joschka Bauer<sup>†</sup> and Thomas Scheibel<sup>\*,†,‡,§,||,⊥,#</sup>

<sup>†</sup>Lehrstuhl Biomaterialien, <sup>‡</sup>Forschungszentrum für Bio-Makromoleküle (BIOMac), <sup>§</sup>Bayreuther Zentrum für Kolloide und Grenzflächen (BZKG), <sup>||</sup>Bayreuther Materialzentrum (BayMat), <sup>⊥</sup>Bayreuther Zentrum für Molekulare Biowissenschaften (BZMB), and <sup>#</sup>Bayrisches Polymerinstitut (BPI), Universität Bayreuth, 95440 Bayreuth, Germany

## Supporting Information

**ABSTRACT:** Spider silk proteins comprise a repetitive core domain with polyalanine and glycine/proline-rich stretches flanked by highly conserved nonrepetitive N- and C-terminal domains. The termini are responsive to assembly triggers, sensing changes in the ionic ( $H^+$ , phosphate) and mechanical (shear stress) environment along the spinning duct. The N-terminal domain dimerizes in a pH-dependent manner induced by protonation of conserved acidic residues. To date, dimerization of N-terminal spider silk domains has been individually investigated in the absence of large core domains. In this work, the impact of an engineered 50 kDa (AQ) core domain was studied on N-terminal dimerization by circular dichroism, fluorescence and absorbance spectroscopy, multiangle light scattering, as well as scanning electron and transmission electron microscopy. Although the core domain showed no apparent influence on the dimerization behavior of the N-terminal domain, the N-terminal domain in contrast influenced the behavior of the core domain: the monomeric state enhanced (AQ)'s solubility, and dimer formation triggered self-assembly. The monomer–dimer equilibrium was influenced by using several previously established mutants, confirming these results. This work thereby provides molecular insights into how key residues of the N-terminal domain control the dimerization-mediated transformation of soluble spidroins into fibrillary assemblies.



## INTRODUCTION

Major ampullate spider silk, also known as dragline silk, is a natural high-performance protein fiber mainly composed of proteins of two major ampullate spidroin classes (MaSp1 and MaSp2), named after the secretory glands in which they are produced. The related MaSp differ in their amino acid compositions, in particular, MaSp2 exhibits a relative high proline content ( $\sim 9\%$ ) in comparison to the nearly proline-free MaSp1 ( $<1\%$ ).<sup>1–4</sup> The unique mechanical properties of dragline silk, especially the high toughness originating from a combination of strength and extensibility, depend on the secondary structure content of the proteins and their self-assembly.<sup>5–7</sup> The amphiphilic spidroins are stored at high concentrations ( $>30\%$  w/v) as a liquid crystalline spinning dope in the lumen of the gland<sup>8,9</sup> and are preassembled into micelle-like structures to prevent premature aggregation.<sup>10,11</sup> The hydrophobic center of the micelles is composed of the intrinsically unstructured core domains of the spidroins, comprising up to 100 repeats of polyalanine and glycine/proline-rich stretches.<sup>1,3</sup> The hydrophilic shell of the micelles is made of explicit, nonrepetitive N- and C-terminal domains both folded into five-helix bundles.<sup>12–14</sup> These terminal domains are highly conserved,<sup>15–17</sup> whereas the type and arrangement of the repetitive core domains differ between spider species and silk types.<sup>5</sup> Although the C-terminal domain of most MaSps exist as

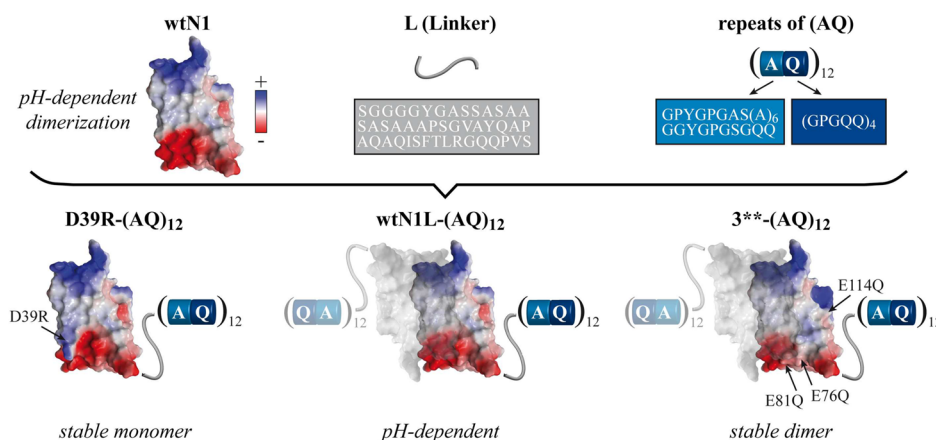
a disulfide-linked, parallel dimer permanently connecting two spidroins,<sup>13</sup> the N-terminal domain forms stable monomers at the storage conditions, neutral pH and relatively high  $Na^+$  and  $Cl^-$  concentrations, in the glands reservoir.<sup>12,18–20</sup> Both termini increase the spidroins' solubility during storage.<sup>12,13</sup> However, the termini also trigger self-assembly by aligning the repetitive modules into  $\beta$ -sheet-rich structures due to environmental changes in the spinning duct,<sup>21</sup> including the emergence of shear stress,<sup>9</sup> a decrease in  $Na^+$  and  $Cl^-$  concentration accompanied by an increase in  $K^+$  and phosphate ( $P_i$ ) concentration,<sup>22</sup> and lowering of the pH from 7.2 to at least 5.7.<sup>23</sup> The addition of  $P_i$  strengthens the intermolecular interactions between the hydrophobic polyalanine stretches of the core domains.<sup>24,25</sup> Acidification-induced protonation of residue E114 stabilizes the dimeric state of the N-terminal domain, physically cross-linking the spidroins into an interconnected network. Simultaneously occurring with dimer formation is the protonation of clustered acidic residues (E76 and E81, whereas D39 tends to lose its proton during dimerization<sup>26</sup>), which independently triggers a further conformational change of the N-terminal domain, thereby

Received: May 11, 2017

Revised: June 22, 2017

Published: June 26, 2017





**Figure 1.** Scheme of studied constructs. Wild-type N-terminal domain of *Latrodectus hesperus* (wtN1) was fused with a short linker (L), designed based on the natural linker sequence, to the engineered core domain (AQ)<sub>12</sub> based on the consensus sequence of ADF3. wtN1L-(AQ)<sub>12</sub> forms monomers at neutral pH and dimers at acidic conditions. Variant D39R-(AQ)<sub>12</sub> is a mutant stabilizing the monomer, and a variant with neutralized conserved acidic amino acid residues stabilizes the dimeric state (3\*\*-(AQ)<sub>12</sub> = E76Q-E81Q-E114Q). The electrostatic surface potential at +5 kT e<sup>-1</sup> (blue) and -5 kT e<sup>-1</sup> of variants of wtN1<sup>18</sup> is shown using PyMOL.<sup>31</sup>

stabilizing the dimer.<sup>19,27,28</sup> The dimerization and conformational change of the wild-type N-terminal domain was not affected by fusion of a small engineered core domain ( $\leq 5$  repetitive sequences).<sup>27,29</sup> The identified stepwise dimerization of individual N-terminal domains was suggested to be a prerequisite for correct spidroin assembly;<sup>19</sup> however, the impact of larger core domains ( $>10$  repetitive sequences) on that behavior and vice versa was not investigated on the molecular level so far. In this work, recombinant spidroins were engineered comprising variants of the N-terminal domain of *Latrodectus hesperus* (wtN1) and a 50 kDa core domain (AQ)<sub>12</sub> designed based on the consensus sequence of *A. diadematus* fibroin 3 (ADF3),<sup>24</sup> and the molecular interplay between both domains was investigated.

## MATERIALS AND METHODS

**Protein Production.** The gene encoding the N-terminal domain of *Latrodectus hesperus* major ampullate spidroin 1 (wtN1) was linked via a short linker (L) to the engineered major dragline silk core sequence (AQ)<sub>12</sub> (Table S1) derived from *Araneus diadematus* fibroin 3 (ADF3) as previously published.<sup>24</sup> The gene encoding wtN1L-(AQ)<sub>12</sub> was cloned into pET28a containing the sequence of a His<sub>6</sub>-SUMO-tag (Novagen). Mutations were introduced using a QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies). *Escherichia coli* (*E. coli*) strain BL21 (DE3) (Agilent Technologies) was transformed and grown at 30 °C to an OD<sub>600</sub> of 60 in Luria-Bertani medium supplemented with Kanamycin. Isopropyl  $\beta$ -D-1-thiogalactopyranoside was added to a final concentration of 1 mM and further incubated for 3 h. Cells were harvested by centrifugation and lysed using a high pressure homogenizer. After centrifugation at 27,000g for 1 h at 4 °C, the clear lysate was loaded onto a nickel-NTA column (HisTrap FF, GE Healthcare). The His<sub>6</sub>-SUMO-tag was proteolytically cleaved by incubation with SUMO protease. Spidroins were precipitated in the presence of 20% (w/v) ammonium sulfate at 4 °C and harvested by centrifugation at 14,000 g for 45 min. (AQ)<sub>24</sub> was used as a control and cloned, expressed, and purified as previously described.<sup>24</sup>

**Sample Preparation.** Lyophilized proteins were dissolved in 6 M guanidinium thiocyanate and dialyzed at room temperature (RT) against 3  $\times$  5 L of 20 mM Tris/HCl, pH 7.6, using dialysis membranes with a molecular weight cutoff of 12–14 kDa (SERVAPOR, SERVA). Protein aggregates were separated by centrifugation at 130,000g for 30 min at RT. Immediately before analysis, protein solutions were diluted to the desired concentration by addition of 50 mM sodium phosphate buffer (NaP<sub>i</sub>, pH 7.2 or 5.5  $\pm$  300 mM NaCl).

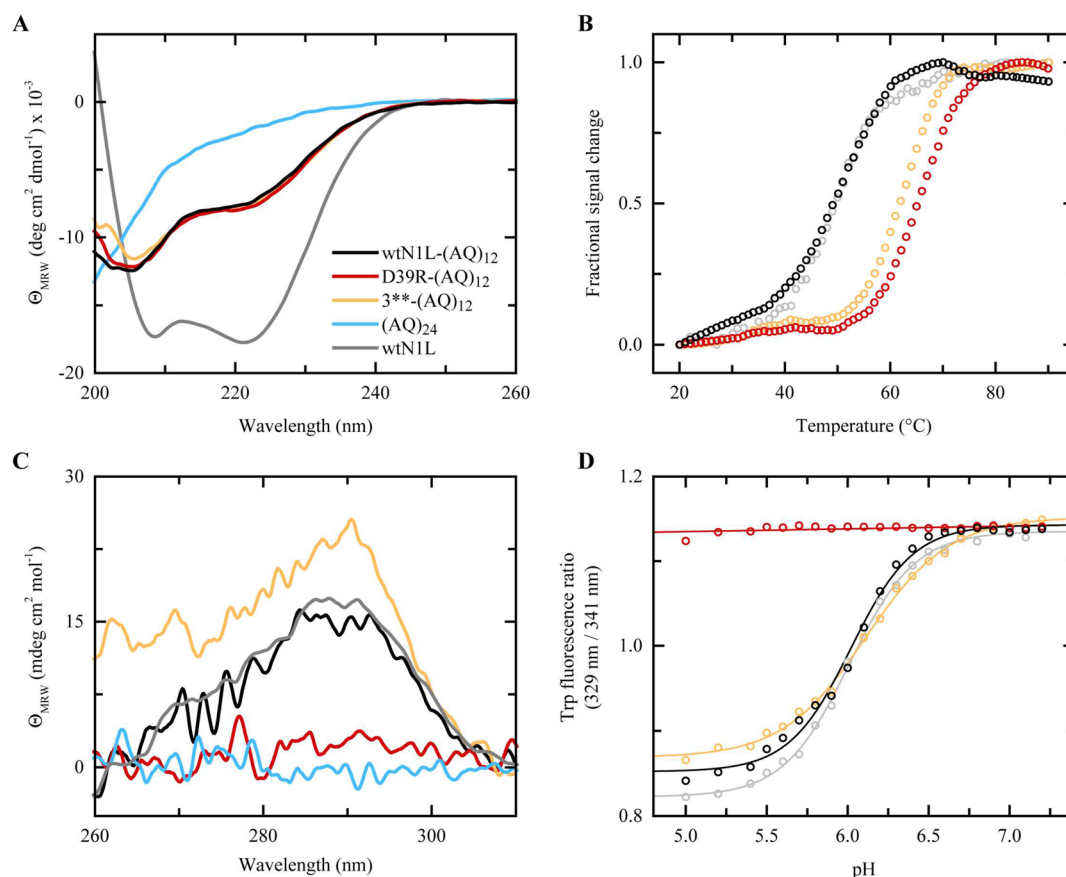
**Circular Dichroism Spectroscopy.** Far- and near-UV circular dichroism (CD) spectra were recorded in cuvettes with 0.1 and 0.5 cm path lengths, respectively, using a Jasco J-715 spectropolarimeter. The scan speed was 50 nm min<sup>-1</sup>, response time 1 s, acquisition interval 0.1 nm, and bandwidth 1 nm. The spectra were calculated as an average of three measurements and were subsequently smoothed by applying a Savitzky–Golay filter. Temperature scans were monitored at a wavelength of 222 nm by elevating the temperature at 1 °C min<sup>-1</sup>.

**Fluorescence Spectroscopy.** Tryptophan fluorescence was monitored with an excitation wavelength of 295 nm, and emission spectra were taken between 300 and 400 nm employing a Jasco FP-6500 spectrofluorometer. The pH titration was conducted by successively adding 0.5 M H<sub>3</sub>PO<sub>4</sub>.

**Size Exclusion Chromatography Multiangle Light Scattering (SEC-MALS).** Mixtures of 2.0 mg mL<sup>-1</sup> of protein were preprepared by size-exclusion chromatography (SEC) (Superdex 200 10/300 GL, GE Healthcare). An Agilent 1100 HPLC system was employed using 50 mM NaP<sub>i</sub> (pH 7.2 or 5.5  $\pm$  300 mM NaCl) at a flow rate of 0.7 mL min<sup>-1</sup>, and 100  $\mu$ L of the protein mixture was injected at a flow rate of 0.2 mL min<sup>-1</sup>. The molecular weight (MW) was subsequently analyzed by multiangle light scattering (MALS) and quasi-elastic light scattering (QELS, WYATT). For time-dependent measurements, the samples were directly injected into a microcuvette, and batch MALS was measured. Prior to this, samples were filtered using a 0.02  $\mu$ m syringe filter, and then aggregation of 0.5 mg mL<sup>-1</sup> protein was initiated by adding KP<sub>i</sub> to a final concentration of 200 mM. The data were processed using ASTRA software (WYATT).

**Transmission Electron Microscopy (TEM).** For analysis of the aggregate morphology, 0.5 mg mL<sup>-1</sup> of each spidroin solution was incubated overnight in 200 mM KP<sub>i</sub> (pH 7.2). Approximately 10  $\mu$ L of the preaggregated protein suspension was spotted on supports (Pioloform 100 mesh copper grids (Plano GmbH)) washed with 10  $\mu$ L of ddH<sub>2</sub>O (Millipore) and negatively stained upon incubation with 5  $\mu$ L of 2% uranyl acetate solution for 2 min. A JEM-2100 transmission electron microscope (JEOL) was used at 80 kV, and images were taken using a 4000  $\times$  4000 charge-coupled device camera (UltraScan 4000; Gatan) and Gatan Digital Micrograph software (version 1.83.842).

**Fourier-Transform Infrared Spectroscopy (FTIR).** FTIR spectroscopy was performed in absorbance mode using a Bruker Tensor 27 spectrometer (Bruker). Aggregation was induced upon addition of 200 mM KP<sub>i</sub> to 10 mg mL<sup>-1</sup> spidroin solutions and incubated overnight. The aggregates were placed on an ATR crystal, and FTIR spectra were taken comprising a 120 scan interferogram with a 4 cm<sup>-1</sup> resolution. The amide I region (1720–1590 cm<sup>-1</sup>) was baseline corrected and smoothed using OPUS software (version 6.5).



**Figure 2.** Structural analysis of the employed constructs. The color coding is indicated in A. (A) Far-UV CD spectra of 0.2 mg mL<sup>-1</sup> of protein were recorded at pH 7.2. (B) Unfolding was analyzed as a function of temperature using CD ellipticity at 222 nm (see also Figure S1). (C) Differential spectra of near-UV CD signal of 2.0 mg mL<sup>-1</sup> of protein between that at pH 7.2 and at pH 5.5, indicating structural changes (see also Figure S2). (D) Ratio of fluorescence at 329 and 341 nm indicating the proportion between conformation I and II (see also Figure S3). The lines represent fits based on a two-state model.

**SDS Stability Analysis.** Aggregation of 2.0 mg mL<sup>-1</sup> of protein was induced by addition of 200 mM KP<sub>i</sub> and incubation overnight. SDS-PAGE sample buffer containing 1% (w/v) SDS was added; samples were heated for 10 min at 95 °C, and 1.5 µg of the protein samples were loaded onto a precast 4–12% acrylamide gradient gel (SERVAGel TG PRiME, SERVA). After SDS-PAGE, the gel was silver stained.

**Photography.** Samples were incubated overnight at a protein concentration of 5.0 mg mL<sup>-1</sup> while rotating at 25 rpm. The resulting aggregates were placed in glass tubes and photographed with a D5100 camera (Nikon).

**Scanning Electron Microscopy (SEM).** Shear-induced aggregates were spotted onto a silica wafer, and samples were sputtered with 1.3 nm platinum using a HR208 sputter coater (Cressington Scientific Instruments). SEM was performed with a Sigma 300VP microscope (Zeiss) using a Se2 detector and an accelerating voltage of 3 kV.

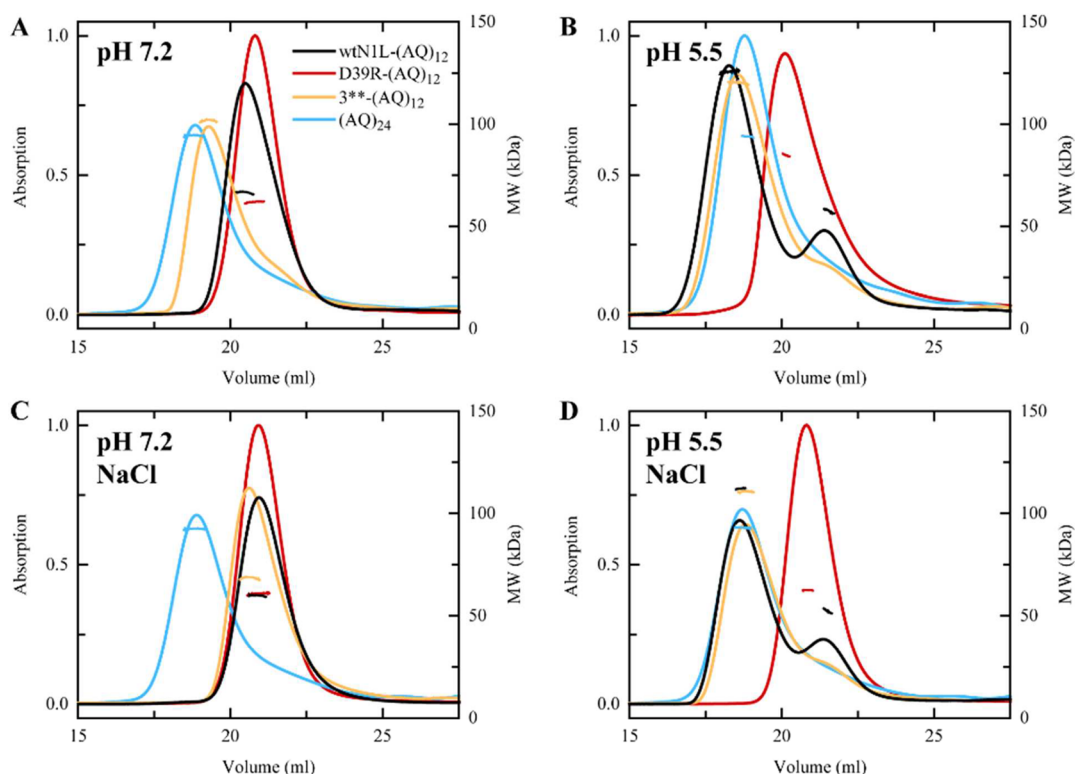
## RESULTS

**Designing Variants of wtN1L-(AQ)<sub>12</sub> with Modified Charge Distribution within the wtN1 Domain.** In the natural spinning process, the transformation of soluble major ampullate spidroins (MaSp) into a tough fiber is triggered by dimerization and conformational switching of their helical amino-terminal domains. For the N-terminal domain of *Latrodectus hesperus* MaSp1 (wtN1), it was previously shown that protonation of a conserved cluster of acidic residues (D39, E76, and E81) leads to the conversion from conformation I (as intended to be in the gland) to conformation II (as speculated

to trigger assembly), and residue E114 mediates the dimer formation.<sup>19</sup> Reversing and neutralizing the charges of these key residues enabled the production of N1 variants that are pH-independently stabilized in the monomeric (D39R) or dimeric state (4\* = D39N-E76Q-E81Q-E114Q). D39 has a strong tendency to lose its proton upon dimerization because it stabilizes the dimer by forming an intermolecular salt bridge,<sup>26,30</sup> and therefore, the modification was excluded here resulting in variant 3\*\* (E76Q-E81Q-E114Q). An analogous variant was already designed for two homologue N-terminal domains.<sup>15,27</sup>

In this work, variants of wtN1 were fused via a short linker (L) with a previously established engineered domain (AQ)<sub>12</sub> based on the core consensus sequence of *Araneus diadematus* fibroin 3 (ADF3) (Figure 1).<sup>24</sup> The resulting wtN1L-(AQ)<sub>12</sub>, D39R-(AQ)<sub>12</sub>, and 3\*\*-(AQ)<sub>12</sub> were characterized to expand the knowledge on how the dimerization of the N-terminal domain controls the assembly of recombinant spidroins and how the intrinsically disordered core domains influence the dimerization of the N-terminal domain in solution.

**Structural Analysis and Thermal Stability of wtN1L-(AQ)<sub>12</sub>.** The secondary structure of wtN1L-(AQ)<sub>12</sub> and variants thereof were investigated by performing far-ultraviolet circular dichroism (UV-CD) spectroscopy (Figure 2A). The core domain, comprising repeats of the (AQ) module, exhibited a random coil (intrinsically disordered) structure, and wtN1L folded into a helical conformation resembling the five-helix



**Figure 3.** Analysis of monomer–dimer distribution of wild-type and variants of N1L-(AQ)<sub>12</sub>. The color coding is indicated in A. (A–D) The SEC profile of 2.0 mg mL<sup>−1</sup> of protein is illustrated by the normalized UV signal. The molecular weight was determined at pH 7.2 (A, C) and pH 5.5 (B, D) in the absence (A, B) and presence of 300 mM NaCl (C, D) using MALS (see also Figure S4).

bundle, as detected without linker.<sup>14,19</sup> In the fusion construct wtN1L-(AQ)<sub>12</sub>, the structures of the core domain (AQ)<sub>12</sub> and the terminal domain N1 were nearly unaffected upon pH change or NaCl addition (Figure S1). Mutating single (D39R-(AQ)<sub>12</sub>) or multiple acidic amino acids (3\*\*-(AQ)<sub>12</sub>) in the N-terminal domain did not alter the secondary structure as seen previously in the case of the individual N1 domain.<sup>19</sup> However, modifying the charge distribution increased the thermal stability of both variants at neutral pH (Figure 2B, Table S2). Acidification increased the  $T_m$  of wtN1L by 16.1 °C and that of wtN1L-(AQ)<sub>12</sub> by 13 °C, indicating formation of dimers as well as reorganization of the tertiary structure (Figure S2, Table S2). The addition of salt significantly stabilized the secondary structure of D39R-(AQ)<sub>12</sub>, thereby diminishing the effects of pH (Figure S1B, D, F, Table S2). 3\*\*-(AQ)<sub>12</sub> was only weakly affected by acidification ( $\Delta T_m = 1$  °C), confirming the almost pH-independent stability of the variant.<sup>19</sup>

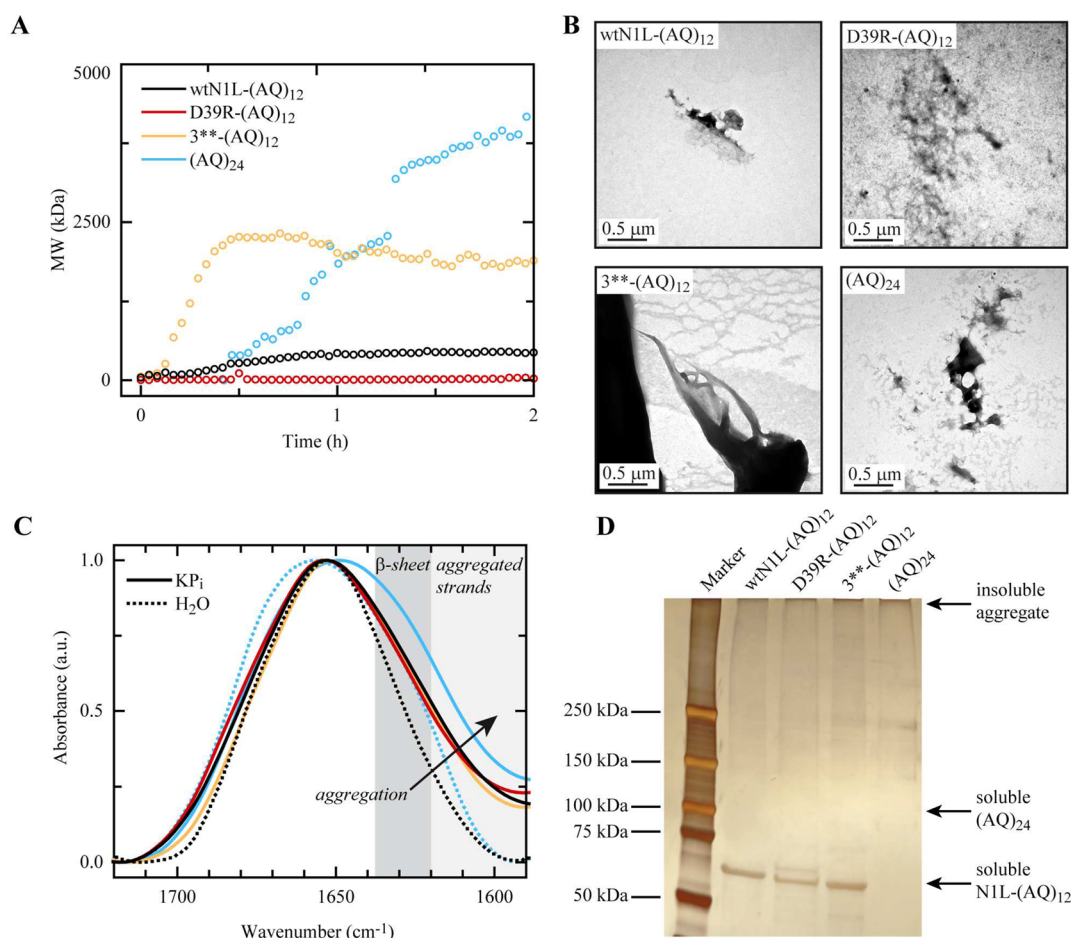
Because the  $T_m$  but not the helical folding of wtN1L-(AQ)<sub>12</sub> was affected by the pH change (Table S2, Figure S1), the wtN1 moiety likely changes its tertiary and/or quaternary structure between pH 7.2 and 5.5 similarly to the already characterized individual domain.<sup>19</sup> Near-UV CD confirmed that wtN1L-(AQ)<sub>12</sub> reorganized its tertiary structure in this pH regime (Figure 2C, Figure S2). The fluorescence of the single tryptophan residue (W9) of wtN1L-(AQ)<sub>12</sub> red-shifted with a transition point at pH 6.0, signaling the conversion from conformation I to the thermally more stable conformation II (Figure 2D, Figure S3). The pH-induced conformational change was only modestly affected in variant 3\*\*-(AQ)<sub>12</sub> but completely suppressed in D39R-(AQ)<sub>12</sub> obviously being restricted to conformation I (Figure 2C and D, Figures S2 and S3). The near-UV CD and fluorescence spectra of the

unstructured (AQ) modules have shown that the pH shift had expectedly no significant influence on the core domain (Figure 2, Figures S2 and S3), which is not surprising because (AQ)<sub>12</sub> does not contain any titratable acidic amino acid residues.

**pH-Dependent Monomer–Dimer Equilibrium of wtN1L-(AQ)<sub>12</sub> and Variants Thereof.** Multiangle light scattering analysis was used to investigate the monomer–dimer equilibrium of recombinant spidroins. Independent of the presence of the core domains, the monomeric state of either wtN1<sup>19</sup> or wtN1L-(AQ)<sub>12</sub> (average MW of 60 kDa) was stabilized at neutral pH and in the presence of NaCl, whereas the dimeric state of wtN1L-(AQ)<sub>12</sub> (126 kDa) was supported upon acidification (Figure 3). D39R-(AQ)<sub>12</sub> was considerably stabilized in the monomeric state at both neutral and acidic pH. Similar effects were reported for the individual domain,<sup>19</sup> and the extension with (AQ)<sub>12</sub> apparently did not affect the quaternary structure. In contrast, variant 3\*\*-(AQ)<sub>12</sub> formed nearly pH-independent dimers (102 kDa at pH 7.2) that were only destabilized in the presence of 300 mM NaCl. Neutralization of charged acidic residues energetically favored the dimer binding,<sup>19</sup> and there was apparently no effect on this process by fusion with the core domain. As expected, the MW (94 kDa) of the core domain (AQ)<sub>24</sub> without any N-terminal domain was not affected by pH or salt.

**Impact of Monomer–Dimer Equilibrium of N1 on Spidroin Assembly.** The hypothesis that the dimerization of wtN1 controls spidroin self-assembly emerged from a study on a mini-spidroin including a homologous unmodified N-terminal domain;<sup>12</sup> however, it has not been examined on the molecular level yet how the assembly is effected by wtN1's key residues controlling the dimerization and structural conversion. Therefore, the assembly behavior of stable monomers/dimers



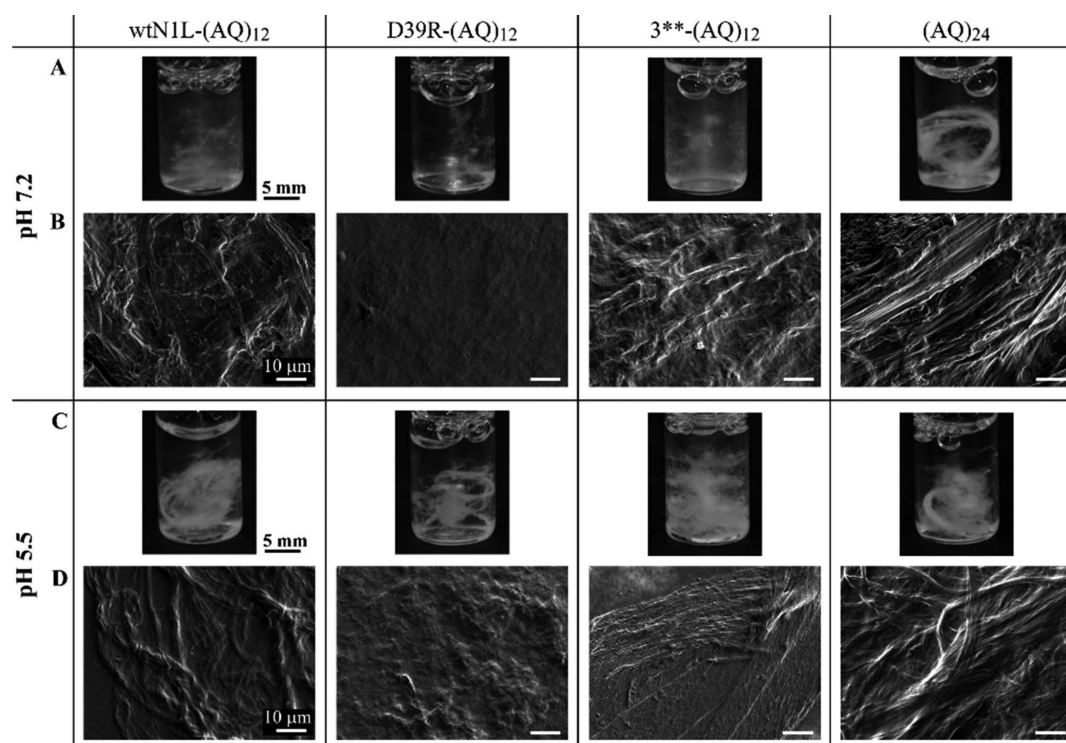


**Figure 4.** Aggregation of wtN1L-(AQ)<sub>12</sub> and variants thereof was initiated upon addition of 200 mM KP<sub>i</sub> (pH 7.2). The color coding is indicated in A. (A) Time-dependent MW changes at 0.5 mg mL<sup>-1</sup> of protein as determined by MALS. (B) TEM images of protein assembled at 0.5 mg mL<sup>-1</sup> in the presence of 200 mM KP<sub>i</sub> incubated for 6 h. (C) Fourier transformed infrared (FTIR) spectra of 10 mg mL<sup>-1</sup> soluble (dotted line, exemplary shown for wtN1L-(AQ)<sub>12</sub> and (AQ)<sub>24</sub>) and aggregated (solid line) protein. Aggregation affected the spectrum, and the affected spectral region is indicated by an arrow. Characteristic wave numbers of  $\beta$ -sheet (dark gray) and aggregate (light gray) structures are highlighted. (D) SDS solubility of aggregates as observed by PAGE. Protein aggregation was induced upon addition of 200 mM KP<sub>i</sub> and incubation overnight. 1% SDS was added, and the samples were boiled for 10 min.

of wtN1L-(AQ)<sub>12</sub> was studied. The assembly of (AQ) modules was induced by the addition of 200 mM potassium phosphate (KP<sub>i</sub>), mimicking the solvent conditions of the spinning duct that was previously used to induce the assembly of a related recombinant spidroin.<sup>9,22,32</sup> (AQ)<sub>24</sub>, acting as a control with almost identical MW but without the wtN1 domain, continuously aggregated into agglomerates with high MW as investigated by MALS (>4,000 kDa, root-mean-square radius of 89 nm) (Figure 4A). Microscopic investigation by transmission electron microscopy (TEM) and bright field microscopy revealed heterogeneous aggregates of (AQ)<sub>24</sub> with a few hundred nm to several μm in diameter (Figure 4B, Figure S5). Fourier-transform infrared spectroscopy showed an IR absorbance maximum at 1655 cm<sup>-1</sup> for soluble (AQ)<sub>24</sub> being characteristic for random coils (Figure 4C).<sup>33</sup> KP<sub>i</sub>-induced aggregation shifted the absorbance maximum (AQ)<sub>24</sub> to a lower wavenumber indicating the formation of  $\beta$ -sheets and aggregated  $\beta$ -strands. The high  $\beta$ -sheet content of (AQ)<sub>24</sub> aggregates was confirmed by its resistance against SDS (Figure 4D) and Thioflavine T binding (Figure S6). Interestingly, the tendency of (AQ) modules to aggregate was reduced at low pH and a low protein concentration of 0.5 mg mL<sup>-1</sup> (Figure S7). The pH of the KH<sub>2</sub>PO<sub>4</sub>/K<sub>2</sub>HPO<sub>4</sub> buffer was lowered by

shifting the mixture toward the monopotassium, thereby decreasing the concentration of kosmotropic K<sup>+</sup> ions, which obviously weakened the intermolecular hydrophobic interactions between the poly-Ala motifs of (AQ)<sub>24</sub> correlating with a reduced salting out of the proteins.<sup>32,34</sup>

Promoting one of the apparent natural functions of the N-terminal domain, premature aggregation of the (AQ)<sub>12</sub> domain was suppressed at neutral pH in the fusion construct wtN1L-(AQ)<sub>12</sub>. MALS indicated that wtN1L-(AQ)<sub>12</sub> agglomerated into smaller structures of approximately 450 kDa (rms radius of 70 nm) (Figure 4A). Interestingly, the formation of aggregates was completely suppressed in variant D39R-(AQ)<sub>12</sub> as measured by MALS (rms radius of 59 nm, 60 kDa) (Figure 4A). TEM confirmed the absence of larger aggregates (Figure 4B). In contrast, the dimeric 3\*\*-(AQ)<sub>12</sub> variant accelerated the assembly kinetics and increased the MW of the agglomerates (~2000 kDa) as seen by MALS (Figure 4A). TEM images showed more ordered and partly aligned morphologies of structures in the micrometer scale, which were also observed using brightfield microscopy (Figure 4B, Figure S5). Interestingly, the helical content of wtN1L-(AQ)<sub>12</sub>, D39R-(AQ)<sub>12</sub>, and 3\*\*-(AQ)<sub>12</sub> was not affected upon aggregation (Figure S8). In comparison to aggregated



**Figure 5.** Shear-induced aggregation of variants of wtN1L-(AQ)<sub>12</sub>. The pH of a 2.0 mg mL<sup>-1</sup> protein solution was adjusted to 7.2 (A, B) and 5.5 (C, D). Samples were rotated overnight, and pictures of the resulting aggregates were taken (A, C). The morphology of the aggregates was imaged by scanning electron microscopy (SEM) (B, D).

(AQ)<sub>24</sub>, the aggregates of all three variants with an N-terminal domain similarly exhibited significantly decreased  $\beta$ -sheet content as indicated by a decreased IR absorbance between 1610 and 1637 cm<sup>-1</sup> (Figure 4C, Figure S8), the solubilization of aggregates by SDS (Figure 4D), and a lower affinity toward Thioflavine T (Figure S6). In summary, these findings confirmed that the monomeric state of wtN1 prevented spidroin aggregation even in the presence of KP<sub>i</sub>, typically acting as an assembly trigger. Acidification induced protonation of conserved acidic key residues (E76, E81, E114), stabilizing the dimeric state of wtN1 and thereby mediating the KP<sub>i</sub>-driven self-assembly without affecting the  $\beta$ -sheet formation of the core domain.

Spidroin self-assembly is further triggered by shear stress arising during its passage through the tapered spinning duct and pulling the fiber from the spider's abdomen. Therefore, the impact of shear stress on the different variants was investigated more closely. The soluble state of wtN1L was not affected by shear stress (data not shown). In contrast, (AQ)<sub>24</sub> formed heterogeneous, fiber-like aggregates in the mm scale upon shearing the solution (Figure 5A). The aggregates consisted of bundles of smooth fibrils with diameters in the  $\mu$ m scale as seen in scanning electron microscopy (SEM) images (Figure 5B) and as was previously reported.<sup>13</sup> Proteins incorporating wtN1 were significantly shifted toward the soluble state because fewer as well as smaller aggregates (diameters of a few  $\mu$ m) occurred at neutral pH (Figure 5A, B). Dimerization of wtN1L-(AQ)<sub>12</sub> at pH 5.5, however, supported the aggregation into bundles of fibrils resembling similar morphologies like that produced by (AQ)<sub>24</sub> (Figure 5C, D). Stabilization of D39R-(AQ)<sub>12</sub> in the monomeric state nearly completely suppressed its aggregation at neutral pH, whereas a few heterogeneous aggregates were formed at low pH. Shifting the equilibrium toward the dimeric

state in 3\*\*-(AQ)<sub>12</sub> again clearly supported the agglomeration; however, the aggregates were less elongated and had a rough surface, indicating that a dynamic monomer–dimer equilibrium might be required for the rearrangement of misplaced spidroins during self-assembly. In general, acidification increased the quantity of assembly, length of the assemblies, and organization into fibrils, suggesting that protonation of additional residues, and thus the structural reorganization, might also affect the self-assembly of 3\*\*-(AQ)<sub>12</sub>. Taken together, these findings showed that key residues of wtN1 control the pH-, KP<sub>i</sub>-, and shear-induced assembly of the (AQ)<sub>12</sub> core domain.

## DISCUSSION

Spiders have developed a unique process of spinning soluble major ampullate spidroins (MaSp) within a few milliseconds into a solid silk fiber with remarkable mechanics.<sup>11</sup> Spidroin self-assembly is triggered along the spinning duct by emerging shear forces and increasing concentrations of K<sup>+</sup> and P<sub>i</sub> strengthening intermolecular interactions between residues of the core domains.<sup>9,22</sup>

Prior to spinning, spidroins are stored in the lumen of the silk gland in an aqueous environment at neutral pH.<sup>23</sup> It is supposed that the spidroins are prearranged into micelle-like structures, exposing their hydrophilic terminal domains to the micelle surface to prevent premature aggregation.<sup>10</sup> Characterization of the engineered spidroin N1L-(AQ)<sub>12</sub> and variants thereof confirmed that the nonrepetitive N-terminal domain of *Latrodectus hesperus* MaSp1 (wtN1) suppresses premature aggregation by increasing the solubility of the core domain at conditions found during spidroin storage such as a neutral pH (Figures 4 and 5). While passing through the spinning duct, the liquid dope becomes acidified, which triggers N1 to form antiparallel dimers, i.e., cross-linking the spidroins/micelles into

an interconnected network.<sup>12,19,23,27</sup> Resembling the features of the individual domain,<sup>19</sup> the property of wtN1 to form pH-dependent dimers was maintained in the engineered construct wtN1L-(AQ)<sub>12</sub> (Figure 3). Whereas the C-terminal domain was shown to control the shear-induced aggregation of (AQ) modules into fibril-like structures,<sup>13</sup> the effect of shear stress on spidroin constructs comprising the N-terminal domain had yet to be examined. Here, it was shown that the pH-dependent monomer–dimer equilibrium of wtN1L-(AQ)<sub>12</sub> significantly affects spidroin self-assembly. Mutating conserved acidic key residues of wtN1, the equilibrium of the monomeric/dimeric state was minimally affected by pH (Figure 3). Reversing the charge of a highly conserved aspartic acid residue (D39R) of wtN1L-(AQ)<sub>12</sub> clearly stabilized the monomeric state (Figure 3), which significantly suppressed the P<sub>i</sub>- and shear-induced aggregation but enhanced the spidroin solubility (Figures 4 and 5). In contrast, neutralization of charges of three conserved acidic amino acid residues favored the dimeric state of wtN1L-(AQ)<sub>12</sub> (3<sup>\*\*-</sup>-(AQ)<sub>12</sub> = E76Q-E81Q-E114Q) (Figure 3). In comparison to the monomeric protein, the dimeric 3<sup>\*\*-</sup>-(AQ)<sub>12</sub> showed a considerably increased self-assembly into fibrillar aggregates (Figures 4 and 5).

Adjusting the monomer–dimer equilibrium through the N-terminal domain significantly affected the P<sub>i</sub>- and shear-induced assembly of recombinant spidroins. This finding reflects the putative natural function of the monomeric N-terminal domain as a “solubilizer” at physiological pH and as an assembly trigger upon dimerization in an acidic environment.

## CONCLUSIONS

In conclusion, it could be shown that the transition of soluble spidroins into solid silk fibers can be controlled by shifting the monomer–dimer equilibrium of their N-terminal domain. At neutral pH, the monomeric state enhances the spidroin's solubility during spidroin storage within the gland's reservoir. A decreasing pH along the spinning duct triggers the protonation of conserved acidic residues stabilizing the dimeric state supporting the P<sub>i</sub>- and shear-induced self-assembly of the core domains into solid morphologies such as fibrils and fibers.

## ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biomac.7b00672.

Extended experimental section, amino acid sequence and characteristics of engineered spidroins, pH- and salt-dependent CD measurements, pH-dependent Trp-fluorescence spectra, pH- and salt-dependent SEC-MALS chromatograms, brightfield microscopy images, Thioflavine T fluorescence spectra, additional aggregation analysis at pH 7.2 and pH 5.5, and pH- and salt-dependent melting temperatures (PDF)

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: thomas.scheibel@bm.uni-bayreuth.de.

### ORCID

Thomas Scheibel: 0000-0002-0457-2423

## Author Contributions

J.B. and T.S., conceptualization; J.B., investigation; J.B., writing (original draft); T.S., writing (review and editing); J.B., visualization; T.S., supervision.

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Dr. Martin Humenik for experimental help, Tamara Aigner for help with TEM, Christian Borkner for help with FTIR, Dr. Hendrik Bargel for help with SEM, and Andreas Schmidt for help with fermentation. Furthermore, we thank Dr. Martin Humenik and Elise DeSimone for proof reading, and Fabian Müller for finalizing the manuscript. Funding was obtained from European Union project ID 123.

## REFERENCES

- (1) Ayoub, N. A.; Garb, J. E.; Tinghitella, R. M.; Collin, M. A.; Hayashi, C. Y. Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. *PLoS One* **2007**, *2* (6), e514.
- (2) Hayashi, C. Y.; Shipley, N. H.; Lewis, R. V. Hypotheses that correlate the sequence, structure, and mechanical properties of spider silk proteins. *Int. J. Biol. Macromol.* **1999**, *24* (2–3), 271–275.
- (3) Hinman, M. B.; Lewis, R. V. Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber. *J. Biol. Chem.* **1992**, *267* (27), 19320–19324.
- (4) Guerette, P. A.; Ginzinger, D. G.; Weber, B. H.; Gosline, J. M. Silk properties determined by gland-specific expression of a spider fibroin gene family. *Science* **1996**, *272* (5258), 112–115.
- (5) Blackledge, T. A.; Hayashi, C. Y. Silken toolkits: biomechanics of silk fibers spun by the orb web spider *Argiope argentata* (Fabricius 1775). *J. Exp. Biol.* **2006**, *209* (13), 2452–2461.
- (6) Madsen, B.; Shao, Z. Z.; Vollrath, F. Variability in the mechanical properties of spider silks on three levels: interspecific, intraspecific and intraindividual. *Int. J. Biol. Macromol.* **1999**, *24* (2–3), 301–306.
- (7) Vollrath, F. Strength and structure of spiders' silks. *Rev. Mol. Biotechnol.* **2000**, *74* (2), 67–83.
- (8) Hijirida, D. H.; Do, K. G.; Michal, C.; Wong, S.; Zax, D.; Jelinski, L. W. 13C NMR of *Nephila clavipes* major ampullate silk gland. *Biophys. J.* **1996**, *71* (6), 3442–3447.
- (9) Chen, X.; Knight, D. P.; Vollrath, F. Rheological characterization of *Nephila* spidroin solution. *Biomacromolecules* **2002**, *3* (4), 644–648.
- (10) Jin, H. J.; Kaplan, D. L. Mechanism of silk processing in insects and spiders. *Nature* **2003**, *424* (6952), 1057–1061.
- (11) Vollrath, F.; Knight, D. P. Liquid crystalline spinning of spider silk. *Nature* **2001**, *410* (6828), 541–548.
- (12) Askarieh, G.; Hedhammar, M.; Nordling, K.; Saenz, A.; Casals, C.; Rising, A.; Johansson, J.; Knight, S. D. Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature* **2010**, *465* (7295), 236–238.
- (13) Hagn, F.; Eisoldt, L.; Hardy, J. G.; Vendrely, C.; Coles, M.; Scheibel, T.; Kessler, H. A conserved spider silk domain acts as a molecular switch that controls fibre assembly. *Nature* **2010**, *465* (7295), 239–242.
- (14) Bauer, J.; Scheibel, T. Conformational stability and interplay of helical N- and C-terminal domains with implications on major ampullate spidroin assembly. *Biomacromolecules* **2017**, *18* (3), 835–845.
- (15) Otikovs, M.; Chen, G.; Nordling, K.; Landreh, M.; Meng, Q.; Jorvall, H.; Kronqvist, N.; Rising, A.; Johansson, J.; Jaudzems, K. Diversified Structural Basis of a Conserved Molecular Mechanism for pH-Dependent Dimerization in Spider Silk N-Terminal Domains. *ChemBioChem* **2015**, *16* (12), 1720–1724.
- (16) Garb, J. E.; Ayoub, N. A.; Hayashi, C. Y. Untangling spider silk evolution with spidroin terminal domains. *BMC Evol. Biol.* **2010**, *10*, 243.



- (17) Bini, E.; Knight, D. P.; Kaplan, D. L. Mapping domain structures in silks from insects and spiders related to protein assembly. *J. Mol. Biol.* **2004**, *335* (1), 27–40.
- (18) Hagn, F.; Thamm, C.; Scheibel, T.; Kessler, H. pH-Dependent Dimerization and Salt-Dependent Stabilization of the N-terminal Domain of Spider Dragline Silk—Implications for Fiber Formation. *Angew. Chem., Int. Ed.* **2011**, *50* (1), 310–313.
- (19) Bauer, J.; Schaal, D.; Eisoldt, L.; Schweimer, K.; Schwarzing, S.; Scheibel, T. Acidic Residues Control the Dimerization of the N-terminal Domain of Black Widow Spiders' Major Ampullate Spidroin 1. *Sci. Rep.* **2016**, *6*, 34442.
- (20) Jaudzems, K.; Askarieh, G.; Landreh, M.; Nordling, K.; Hedhammar, M.; Jornvall, H.; Rising, A.; Knight, S. D.; Johansson, J. pH-dependent dimerization of spider silk N-terminal domain requires relocation of a wedged tryptophan side chain. *J. Mol. Biol.* **2012**, *422* (4), 477–487.
- (21) Eisoldt, L.; Thamm, C.; Scheibel, T. Review the role of terminal domains during storage and assembly of spider silk proteins. *Biopolymers* **2012**, *97* (6), 355–361.
- (22) Knight, D. P.; Vollrath, F. Changes in element composition along the spinning duct in a *Nephila* spider. *Naturwissenschaften* **2001**, *88* (4), 179–82.
- (23) Andersson, M.; Chen, G. F.; Otikovs, M.; Landreh, M.; Nordling, K.; Kronqvist, N.; Westermarck, P.; Jornvall, H.; Knight, S.; Ridderstrale, Y.; Holm, L.; Meng, Q.; Jaudzems, K.; Chesler, M.; Johansson, J.; Rising, A. Carbonic Anhydrase Generates CO<sub>2</sub> and H<sup>+</sup> That Drive Spider Silk Formation Via Opposite Effects on the Terminal Domains. *PLoS Biol.* **2014**, *12* (8), e1001921.
- (24) Huemmerich, D.; Helsen, C. W.; Quedzuweit, S.; Oschmann, J.; Rudolph, R.; Scheibel, T. Primary structure elements of spider dragline silks and their contribution to protein solubility. *Biochemistry* **2004**, *43* (42), 13604–13612.
- (25) Eisoldt, L.; Hardy, J. G.; Heim, M.; Scheibel, T. R. The role of salt and shear on the storage and assembly of spider silk proteins. *J. Struct. Biol.* **2010**, *170* (2), 413–419.
- (26) Barroso da Silva, F. L.; Pasquali, S.; Derreumaux, P.; Dias, L. G. Electrostatics analysis of the mutational and pH effects of the N-terminal domain self-association of the major ampullate spidroin. *Soft Matter* **2016**, *12* (25), 5600–5612.
- (27) Kronqvist, N.; Otikovs, M.; Chmyrov, V.; Chen, G.; Andersson, M.; Nordling, K.; Landreh, M.; Sarr, M.; Jornvall, H.; Wennmalm, S.; Widengren, J.; Meng, Q.; Rising, A.; Otzen, D.; Knight, S. D.; Jaudzems, K.; Johansson, J. Sequential pH-driven dimerization and stabilization of the N-terminal domain enables rapid spider silk formation. *Nat. Commun.* **2014**, *5*, 3254.
- (28) Schwarze, S.; Zwettler, F. U.; Johnson, C. M.; Neuweiler, H. The N-terminal domains of spider silk proteins assemble ultrafast and protected from charge screening. *Nat. Commun.* **2013**, *4*, 2815.
- (29) Landreh, M.; Andersson, M.; Marklund, E. G.; Jia, Q.; Meng, Q.; Johansson, J.; Robinson, C. V.; Rising, A. Mass spectrometry captures structural intermediates in protein fiber self-assembly. *Chem. Commun. (Cambridge, U. K.)* **2017**, *53* (23), 3319–3322.
- (30) Gronau, G.; Qin, Z.; Buehler, M. J. Effect of sodium chloride on the structure and stability of spider silk's N-terminal protein domain. *Biomater. Sci.* **2013**, *1* (3), 276–284.
- (31) Schrodinger, L. PyMOL, molecular graphics system, version 1.8; in 2015.
- (32) Humenik, M.; Magdeburg, M.; Scheibel, T. Influence of repeat numbers on self-assembly rates of repetitive recombinant spider silk proteins. *J. Struct. Biol.* **2014**, *186* (3), 431–437.
- (33) Hu, X.; Kaplan, D.; Cebe, P. Determining beta-sheet crystallinity in fibrous proteins by thermal analysis and infrared- spectroscopy. *Macromolecules* **2006**, *39* (18), 6161–6170.
- (34) Slotta, U. K.; Rammensee, S.; Gorb, S.; Scheibel, T. An engineered spider silk protein forms microspheres. *Angew. Chem., Int. Ed.* **2008**, *47* (24), 4592–4594.

## **Supporting Information**

Dimerization of the conserved N-terminal domain of a spider silk protein controls the self-assembly of the repetitive core domain.

*Joschka Bauer and Thomas Scheibel*



**Amino Acid Sequence:**

**wtN1L**

10	20	30	40	50
SMGQANTPWS	SKANADAFIN	SFISAASNTG	SFSQDQMEDM	SLIGNTLMAA
60	70	80	90	100
MDNMGGRITP	SKLQALDMAF	ASSVAEIAAS	EGGDLGVTTN	AIADALTSAF
110	120	130	140	150
YQTTGVVNSR	FISEIRSLIG	MFAQASANDV	YASAGSGSGG	GGYGASSASA
160	170	180		
ASASAAAPSG	VAYQAPAQAQ	ISFTLRGQQP	VSG	

### **Extended Experimental Section:**

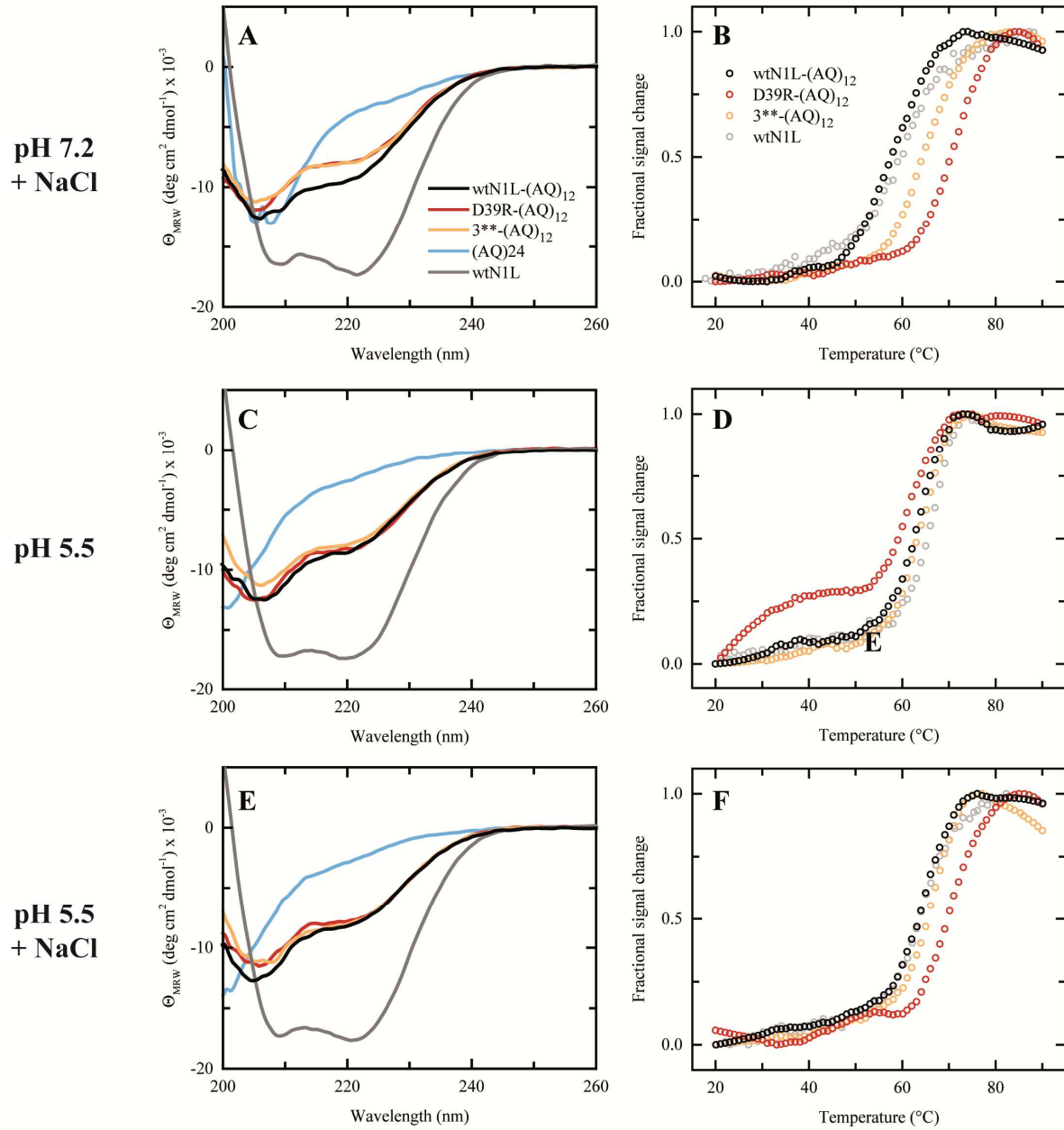
**SDS Stability Analysis.** Aggregation of  $2.0 \text{ mg mL}^{-1}$  protein was induced upon addition of 200 mM potassium phosphate ( $\text{KP}_i$ ) and incubation overnight. SDS-PAGE sample buffer containing 1 % (w/v) SDS was added, samples were heated for 10 min at  $95^\circ\text{C}$ , and  $1.5 \text{ }\mu\text{g}$  of the protein samples were loaded onto a pre-cast 4 - 12 % acrylamide gradient gel (SERVAGel™ TG PRiME™, SERVA). After SDS-PAGE, the gel was silver stained.

**Thioflavine T Binding.** Binding of  $10 \text{ }\mu\text{M}$  Thioflavine T (ThT) to  $40 \text{ }\mu\text{M}$  pre-aggregated protein was measured at an excitation wavelength of 440 nm by recording emission spectra between 450 and 600 nm. Shown ThT fluorescence spectra are an average of three consecutive scans.

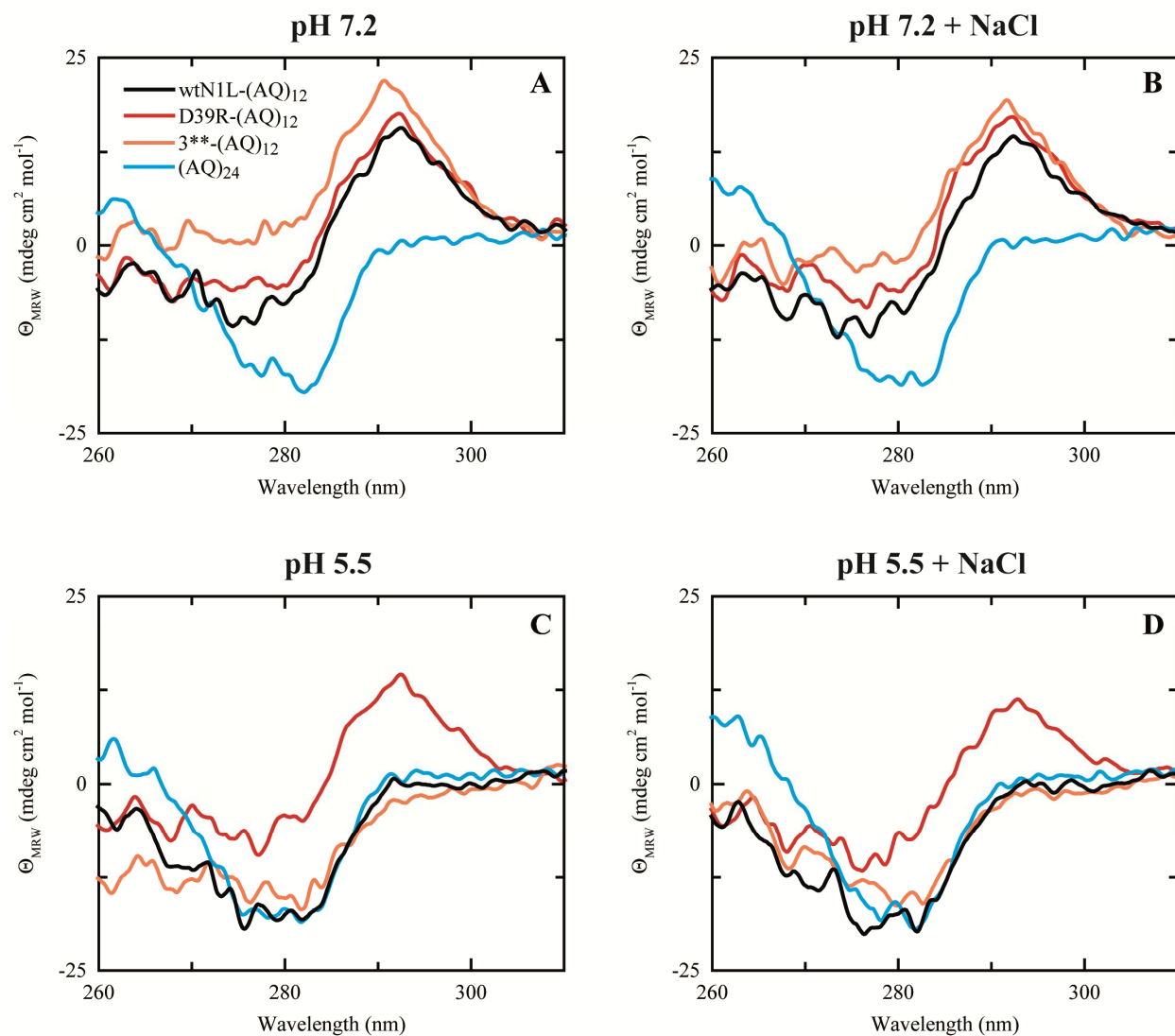
**Optical Microscopy.** The pH of a  $0.5 \text{ mg mL}^{-1}$  protein solution was adjusted to pH 7.2 or 5.5 as indicated, and the aggregation was induced by addition of 200 mM  $\text{KP}_i$  and incubation overnight. The aggregates were transferred onto glass slides for bright-field microscopy (DMI3000B, Leica) using a  $40\times$  objective lens.

**Fourier-Transform Infrared Spectroscopy (FTIR).** The amide I region ( $1590 - 1705 \text{ cm}^{-1}$ ) was analyzed by Fourier self-deconvolution (FSD) using OPUS software (version 6.5). Bands were assigned to secondary structure elements according to Hu et al.<sup>1</sup>

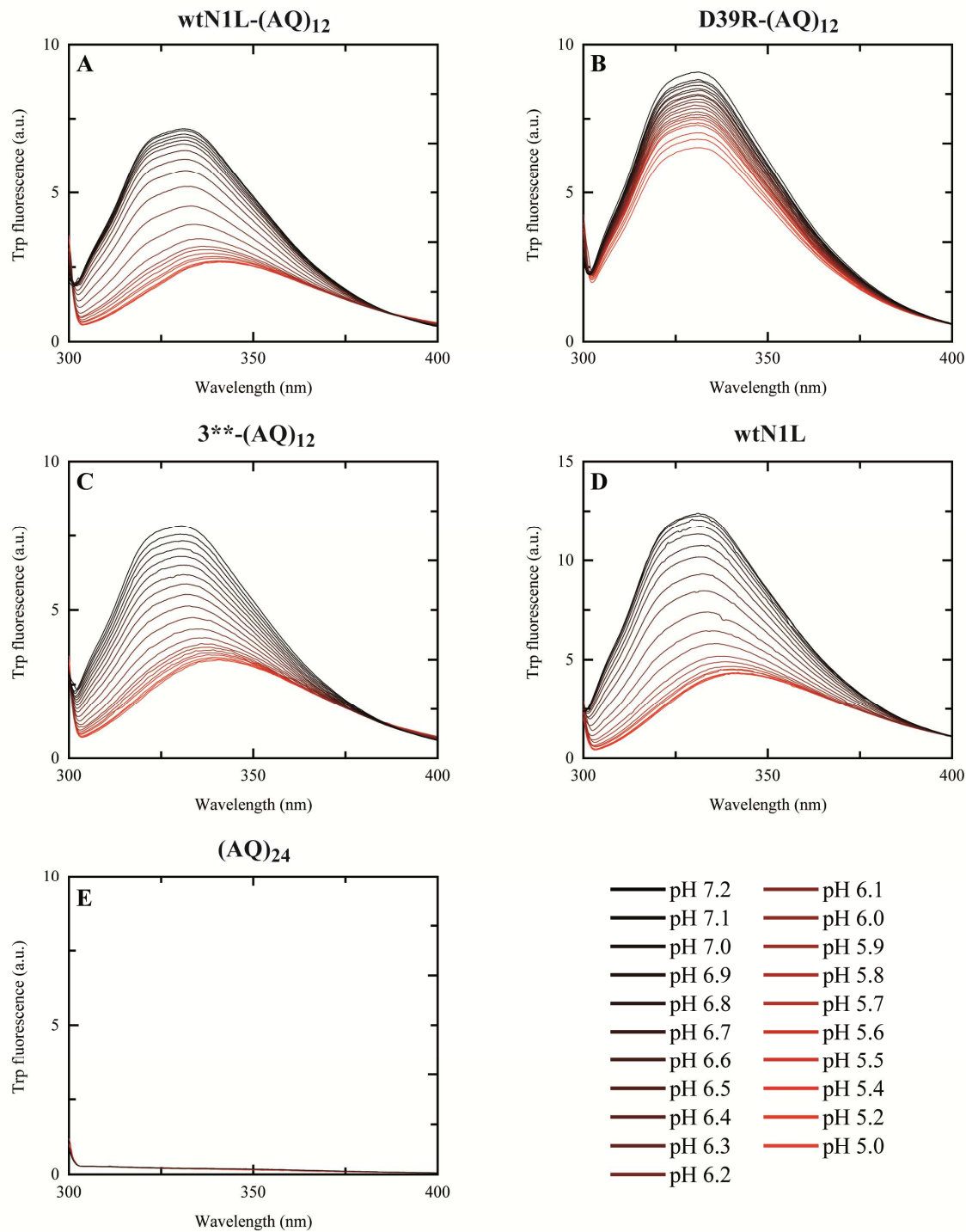
### Additional Figures:



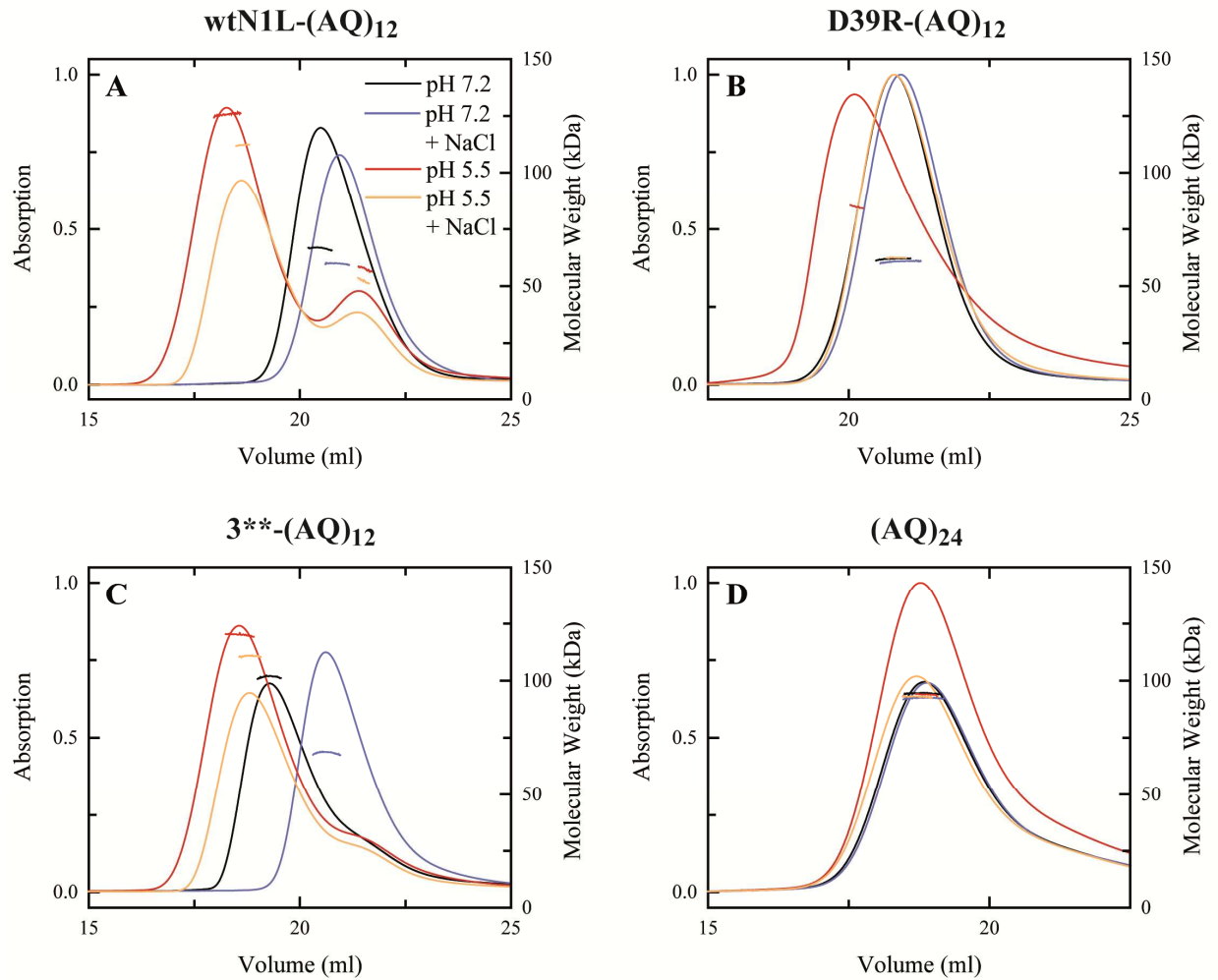
**Figure S1:** Structural analysis (A, C, E) and thermal stability (B, D, F) of all tested constructs at conditions as depicted on the left panel. The residual molar ellipticity of  $0.2 \text{ mg mL}^{-1}$  protein was determined in a pH- and salt-dependent manner. Protein denaturation was analyzed as a function of temperature using CD ellipticity at 222 nm. The intrinsically unstructured (AQ)<sub>24</sub> does not provide a melting curve.



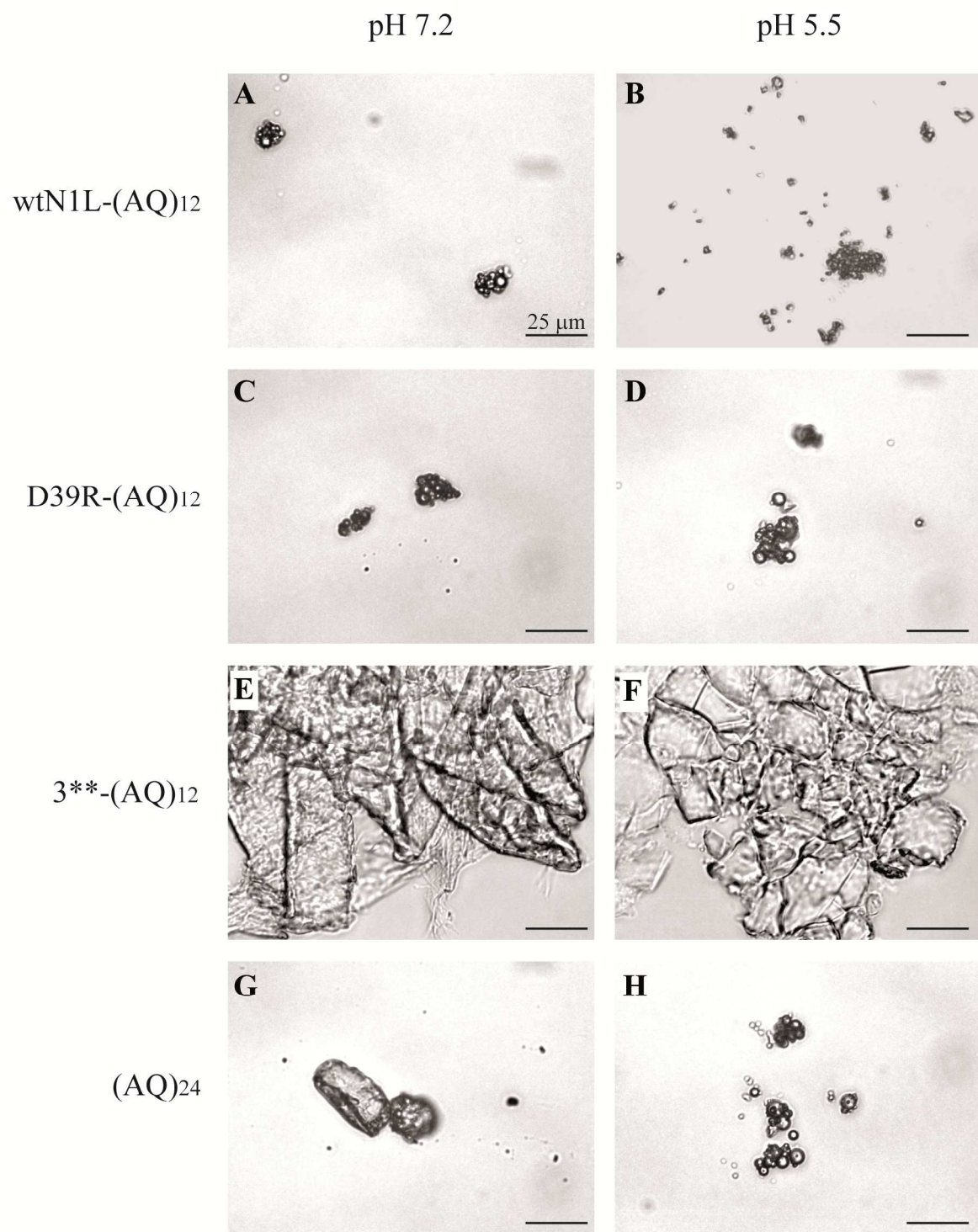
**Figure S2:** Near-UV CD spectra were taken at a protein concentration of  $2.0 \text{ mg mL}^{-1}$  at room temperature.



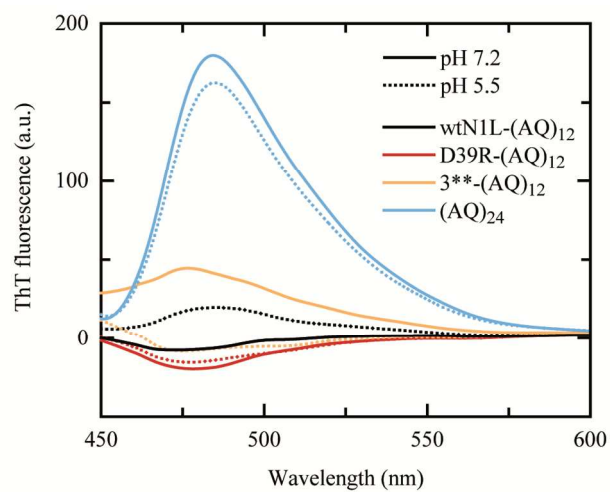
**Figure S3:** Trp fluorescence spectra were recorded at an excitation wavelength of 295 nm and an emission wavelength between 300 and 400 nm. The pH was decreased from 7.2 to 5.0 by successively adding 0.1 M H<sub>3</sub>PO<sub>4</sub> to 0.1 mg mL<sup>-1</sup> protein without exhibiting a detectable effect of dilution.



**Figure S4:** Chromatograms of wildtype and variants of N1L-(AQ)<sub>12</sub> as well as (AQ)<sub>24</sub>. The chromatograms show the UV signal, and the molecular weight of the proteins is indicated. MALS was measured at pH 7.2 (black) as well as pH 5.5 (red) in the absence and presence of 300 mM NaCl (pH 7.2 + NaCl, blue; pH 5.5 + NaCl, orange).

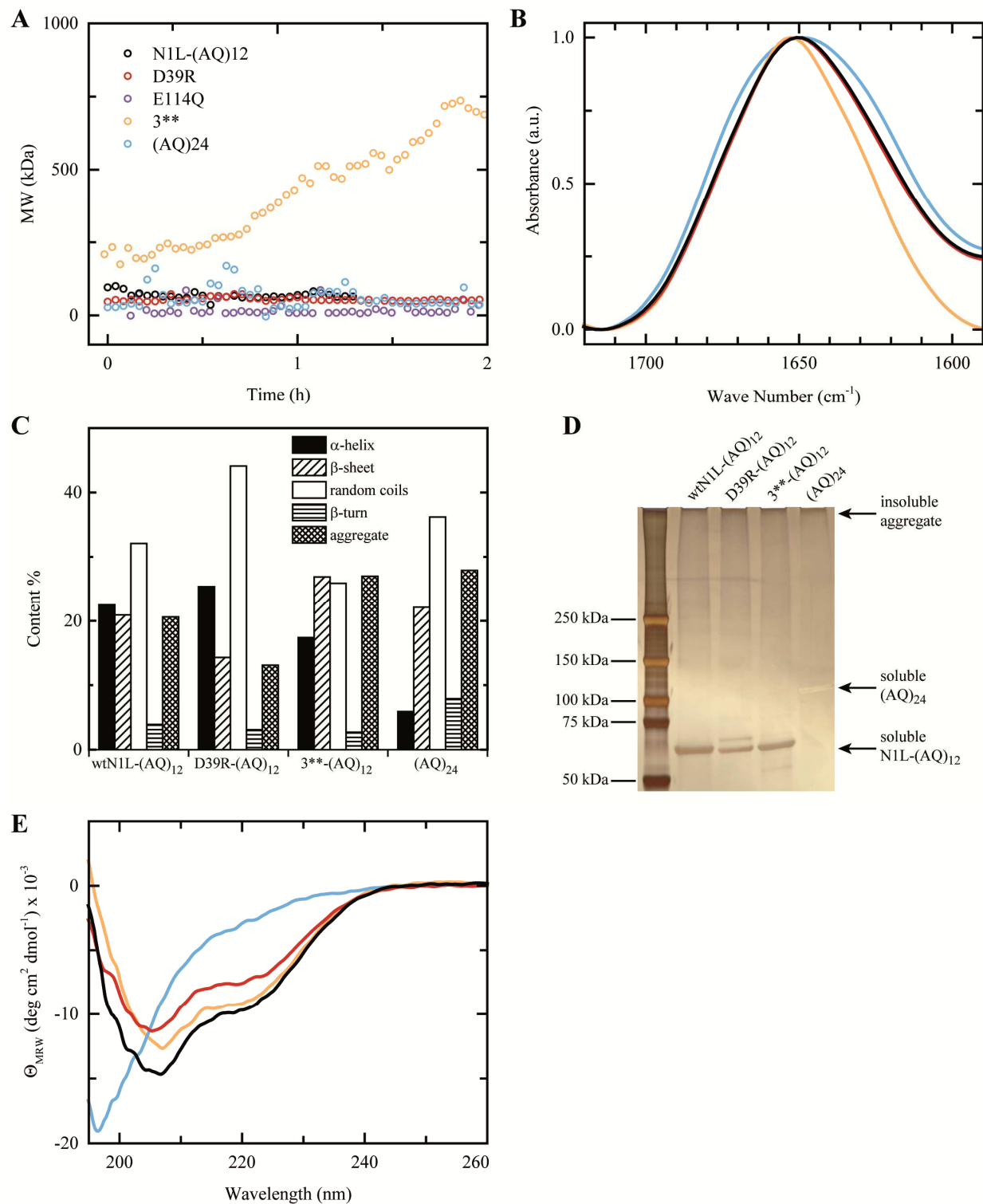


**Figure S5:** Brightfield microscopy images of protein aggregates. The pH of a 0.5 mg mL<sup>-1</sup> protein solution was adjusted to 7.2 or 5.5, and aggregation was induced by addition of 200 mM KP<sub>i</sub> and incubation overnight.



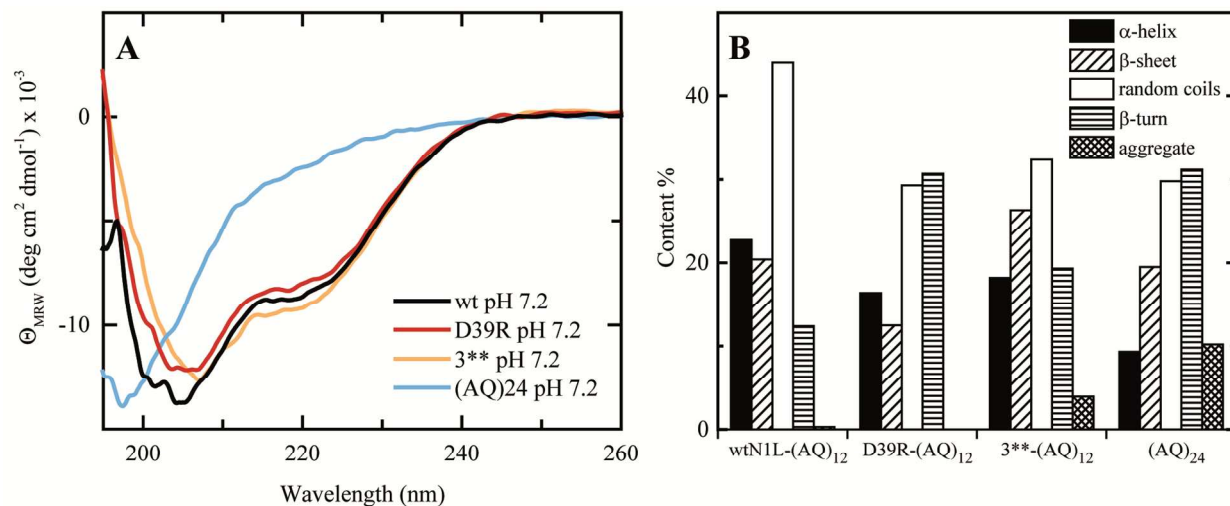
**Figure S6:** Thioflavine T (10  $\mu\text{M}$ ) fluorescence spectra were taken in the presence of 0.5 mg  $\text{mL}^{-1}$  pre-aggregated protein. Protein aggregation was induced by addition of 200 mM  $\text{KP}_i$  and incubation overnight.





**Figure S7:** KP<sub>i</sub>-induced aggregation of wtN1L-(AQ)<sub>12</sub> and variants thereof at pH 5.5. The color coding system is indicated in part A of the figure. **A)** The time-dependent MWs of 0.5 mg mL<sup>-1</sup>

protein as determined by MALS. **B)** Fourier transformed infrared (FTIR) spectra of 10 mg mL<sup>-1</sup> protein aggregates. **C)** Secondary structure distribution (in %) of protein aggregates. Spectra were deconvoluted using Fourier self-deconvolution (FSD) of the amide I band. **D)** SDS solubility of aggregates as observed by PAGE. Protein aggregation was induced by addition of 200 mM KP<sub>i</sub> and incubation overnight. 1 % SDS was added and the samples were boiled for 10 min. **E)** The residual molar ellipticity of 0.2 mg mL<sup>-1</sup> protein aggregates was determined at pH 5.5.



**Figure S8:** KP<sub>i</sub>-induced aggregation of wtN1L-(AQ)<sub>12</sub> and variants thereof at pH 7.2. **A)** The residual molar ellipticity of 0.2 mg mL<sup>-1</sup> protein aggregates was determined at pH 7.2. **B)** Secondary structure distribution (in %) of protein aggregates. Spectra were deconvoluted using Fourier self-deconvolution (FSD) of the amide I band.

**Table S1:** Physico-chemical properties of recombinant spidroins.

	wtN1L	wtN1L-(AQ) <sub>12</sub>	D39R-(AQ) <sub>12</sub>	3**-(AQ) <sub>12</sub>	(AQ) <sub>24</sub>
<b>MW of monomer [kDa]</b>	19.6	66.1	66.1	66.1	94.5
<b>theoretical pI</b>	4.0	4.0	4.1	4.0	4.1
<b>number of negative charges</b>	20	20	19	17	1
<b>number of positive charges</b>	7	7	8	7	2

**Table S2:** pH- and salt-dependent melting temperatures ( $T_m$ ) of N1L-(AQ)<sub>12</sub> variants. Thermal denaturation of 0.2 mg mL<sup>-1</sup> protein was analyzed using CD ellipticity at 222 nm. The min/max values at distinct solvent conditions are highlighted in blue/red.

$T_m$	pH 7.2	pH 7.2 + NaCl	pH 5.5	pH 5.5 + NaCl
wtN1L-(AQ) <sub>12</sub>	49.1 °C	56.6 °C	62.3 °C	63.5 °C
D39R-(AQ) <sub>12</sub>	65.5 °C	70.6 °C	60.4 °C	70.1 °C
3**-(AQ) <sub>12</sub>	61.9 °C	64.6 °C	63.1 °C	64.8 °C
wtN1L	49.4 °C	59.7 °C	65.5 °C	64.1 °C

**Reference**

1. Hu, X.; Kaplan, D.; Cebe, P., Determining beta-sheet crystallinity in fibrous proteins by thermal analysis and infrared- spectroscopy. *Macromolecules* **2006**, 39, (18), 6161-6170.

Teilarbeit IV wurde 2017 bei *Biomacromolecules* veröffentlicht.

Bauer, J. & Scheibel, T. (2017) Conformational Stability and Interplay of Helical N- and C-Terminal Domains with Implications on Major Ampullate Spidroin Assembly. *Biomacromolecules* 18 (3): 835.

Im Folgenden wird die Publikation mit Genehmigung von Copyright © 2017 American Chemical Society reproduziert.

# Conformational Stability and Interplay of Helical N- and C-Terminal Domains with Implications on Major Ampullate Spidroin Assembly

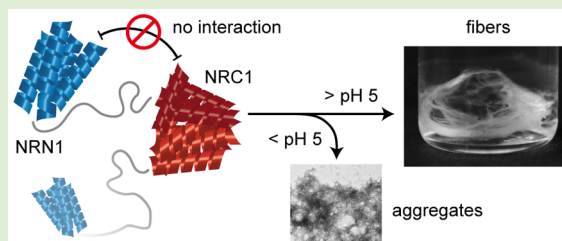
Joschka Bauer<sup>†</sup> and Thomas Scheibel<sup>\*,†,‡,§,||,⊥,#</sup>

<sup>†</sup>Lehrstuhl Biomaterialien, <sup>‡</sup>Forschungszentrum für Bio-Makromoleküle (BIOmac), Bayrisches Geoinstitut, <sup>||</sup>Bayreuther Materialzentrum (BayMat), Fakultät für Ingenieurwissenschaften, <sup>#</sup>Bayrisches Polymerinstitut (BPI), Universität Bayreuth, Universität Bayreuth, Universitätsstr. 30, 95440 Bayreuth, Germany

<sup>§</sup>Bayreuther Zentrum für Kolloide und Grenzflächen (BZKG), <sup>⊥</sup>Bayreuther Zentrum für Molekulare Biowissenschaften (BZMB), Universität Bayreuth, Naturwissenschaften I, 95440 Bayreuth, Germany

## Supporting Information

**ABSTRACT:** Major ampullate spidroin (MaSp) assembly starts in the abdomen of the spider, where spidroins are stored as a liquid dope at a high concentration. The dope is squeezed into the spinning duct, and assembly is finished upon drawing of fibers. Unwanted aggregation of the spidroin solution in the gland is suppressed by prestructuring of the spidroins in micelle-like assemblies, with their hydrophobic stretches being hidden from the solvent and the hydrophilic nonrepetitive amino (NRN) and carboxy (NRC) terminal domains being exposed on the micelle surface. Conversion of the fluid dope into a solid fiber is induced within the spinning duct by acidification and ion exchange (sodium chloride against potassium phosphate), with the impact on the structure of the NRN and NRC domains acting as a regulatory switch for fiber assembly. While NRN dimerizes pH-dependently in an antiparallel fashion (i.e. quaternary structural changes), the tertiary structure of dimeric NRC is changed by shear stress and a drop in pH, inducing the alignment of the intrinsically unstructured core domains accompanied by  $\beta$ -sheet formation of motifs of the core domain. Here, the conformational stability of NRN1 and NRC1 of *Latrodectus hesperus* MaSp1 were studied using independent techniques such as circular dichroism, fluorescence and absorbance spectroscopy, and scanning electron, transmission electron, and atomic force microscopy. In this context, it could be shown that strong, non-natural acidification drives NRC1 to unfold and aggregate into  $\beta$ -sheet-rich structures, preventing recombinant spidroins from assembling into aligned fibrils. Interestingly, NRN1 and NRC1 apparently do not interact with each other, making spidroin assembly easy to control step-by-step and straightforward due to missing unproductive side reactions.



## INTRODUCTION

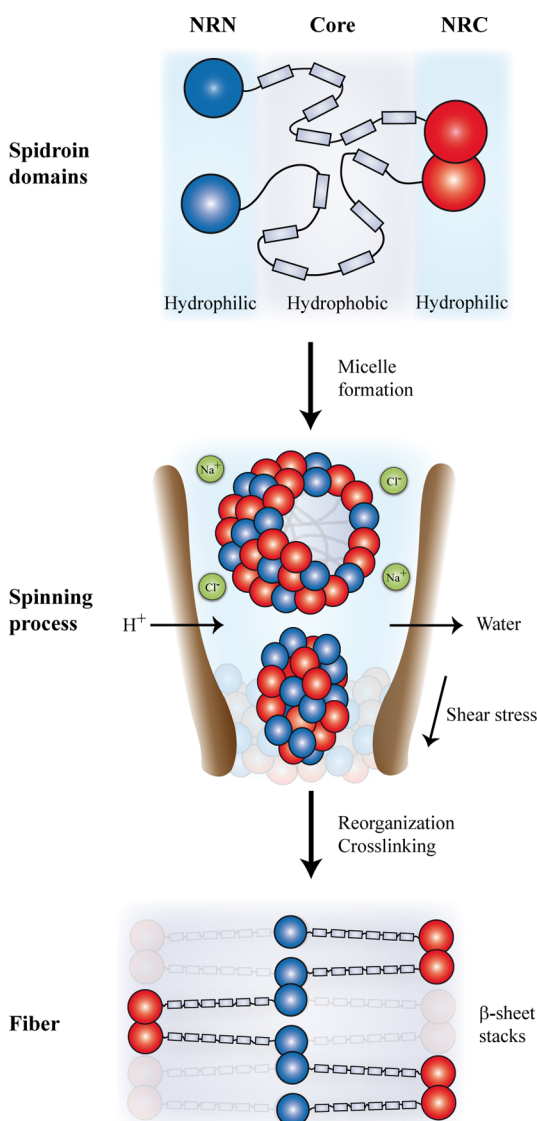
Major ampullate silk, also known as dragline silk, is used as the stabilizing frame and the radii in spider webs and as a lifeline by the spider to escape from predators.<sup>1</sup> Dragline silk is well-adapted to this function due to its combination of strength and extensibility, resulting in a toughness that outperforms most synthetic polymers.<sup>1–3</sup> The mechanical properties can be directly related to the underlying proteins, their sequences, and their assembly mechanism.<sup>4</sup> The proteins comprising the core of the silk fiber are called spidroins and share a common pattern of an intrinsically (in solution) unstructured highly repetitive core domain as well as nonrepetitive amino-terminal (NRN) and carboxy-terminal (NRC) domains both representing five-helix bundles (Figure 1).<sup>5–7</sup> The large core domain (up to 350 kDa) is composed of up to 100 repeating modules, each consisting of 40–200 amino acids with a high content of polyaniline- and glycine-rich motifs.<sup>8–11</sup> During spidroin assembly into fibers, polyaniline stretches convert into  $\beta$ -sheets,<sup>12</sup> which are embedded in random coil/helical structures of glycine-rich areas<sup>13,14</sup> resembling an amorphous matrix.<sup>4,15,16</sup> While the type and arrangement of repetitive motifs vary between different silk types and spiders,<sup>2</sup> the NRN and NRC

domains are highly conserved.<sup>17–22</sup> The terminal domains are important for stabilizing spidroins during storage, and they also control spidroin assembly by alignment of the proteins, which occurs within a few milliseconds after induction by external triggers such as shear stress, acidification, and the replacement of sodium and chloride ions by potassium and phosphate ones.<sup>5,23–25</sup> Acidification and depletion of sodium chloride induce conformational changes in NRN accompanied by antiparallel dimerization, which are important for assembly.<sup>6,22,26,27</sup> Dimeric NRC, in contrast, stabilizes the soluble state of spidroins by inducing a prestructuring in the dope and mediates the shear-induced alignment of repetitive segments into  $\beta$ -sheet structures upon assembly.<sup>7,28</sup> In order to develop spider silk fibers with comparable mechanics to that of natural ones, it is essential to understand the molecular mechanisms occurring during natural assembly.<sup>29</sup> It has been recently suggested that NRC of *Nephila clavipes* MaSp1 and *Araneus ventricosus* MiSp unfolds at low pH followed by amyloid-like

Received: November 19, 2016

Revised: January 25, 2017

Published: January 27, 2017



**Figure 1.** Scheme of the spidroin self-assembly into a fiber. Spidroins are composed of an intrinsically unstructured, highly repetitive core domain as well as globular folded, nonrepetitive amino-terminal (NRN) and carboxy-terminal (NRC) domains. During storage, the spidroins are prearranged into micelles, being solubilized by their hydrophilic termini. Changing chemical and mechanical conditions affect the termini's tertiary and quaternary structures, triggering spidroin reorganization into a  $\beta$ -sheet-rich fiber.

fibril formation.<sup>25,30</sup> In this work, the conformational stabilities of NRN1 and NRC1 of *Latrodectus hesperus* MaSp1 were characterized at different pH values, including nonphysiological conditions, to analyze if amyloid-like fibril formation might be a general feature of terminal domains. Thereby, the impact of strong acidification on the natural role of NRC1 in controlling spidroin self-assembly was investigated. Furthermore, putative interactions between both domains were studied to analyze whether they act in conjunction or separately upon assembly.

## MATERIALS AND METHODS

**Protein Production.** Genes encoding either the carboxy- or the amino-terminal domain of *L. hesperus* major ampullate spidroin 1 were cloned into pET28a containing a His<sub>6</sub>-SUMO-tag (Novagen) sequence. *Escherichia coli* strain BL21 (DE3) (Agilent Technologies) was transformed and grown at 30 °C in Luria–Bertani medium

supplemented with kanamycin to an OD<sub>600</sub> of 60. Isopropyl  $\beta$ -D-1-thiogalactopyranoside was added to a final concentration of 1 mM, followed by further incubation for 3 h. Cells were harvested by centrifugation and disrupted using a high-pressure homogenizer. After centrifugation at 27 000g for 1 h at 4 °C, the cleared lysate was loaded onto a nickel-NTA column (HisTrap FF, GE Healthcare). The His<sub>6</sub>-SUMO tag was cleaved off by incubation with SUMO protease for 1 h at room temperature (RT). Tag and protease were removed using a second nickel-NTA chromatography. Spidroin domains were further purified by size-exclusion chromatography (HiLoad 26/60 Superdex 75 pg, GE Healthcare). Spidroins comprising repetitive modules ((AQ)<sub>24</sub>, (AQ)<sub>12</sub>-NR3, and (AQ)<sub>12</sub>-NRC1) were produced and purified pursuing a previously described strategy including a heat step followed by ammonium sulfate precipitation.<sup>31</sup>

**Sample Preparation.** Lyophilized proteins were solved in 20 mM NaP buffer (pH 7.2) at a concentration of 1 mM, and aggregates were removed by centrifugation at 130 000g for 30 min at RT. Recombinant spidroins (AQ)<sub>24</sub>, (AQ)<sub>12</sub>-NR3, (AQ)<sub>12</sub>-NRC1, and eADF4(C<sub>16</sub>) were solved in 6 M guanidinium thiocyanate and dialyzed against 20 mM Tris buffer (pH 7.2) prior to centrifugation. Immediately before analysis, protein solutions were diluted to the desired concentration by addition of either 20 mM NaP (pH 7.2, pH 6.0) or 20 mM NaOAc (pH 5.0, pH 4.0).

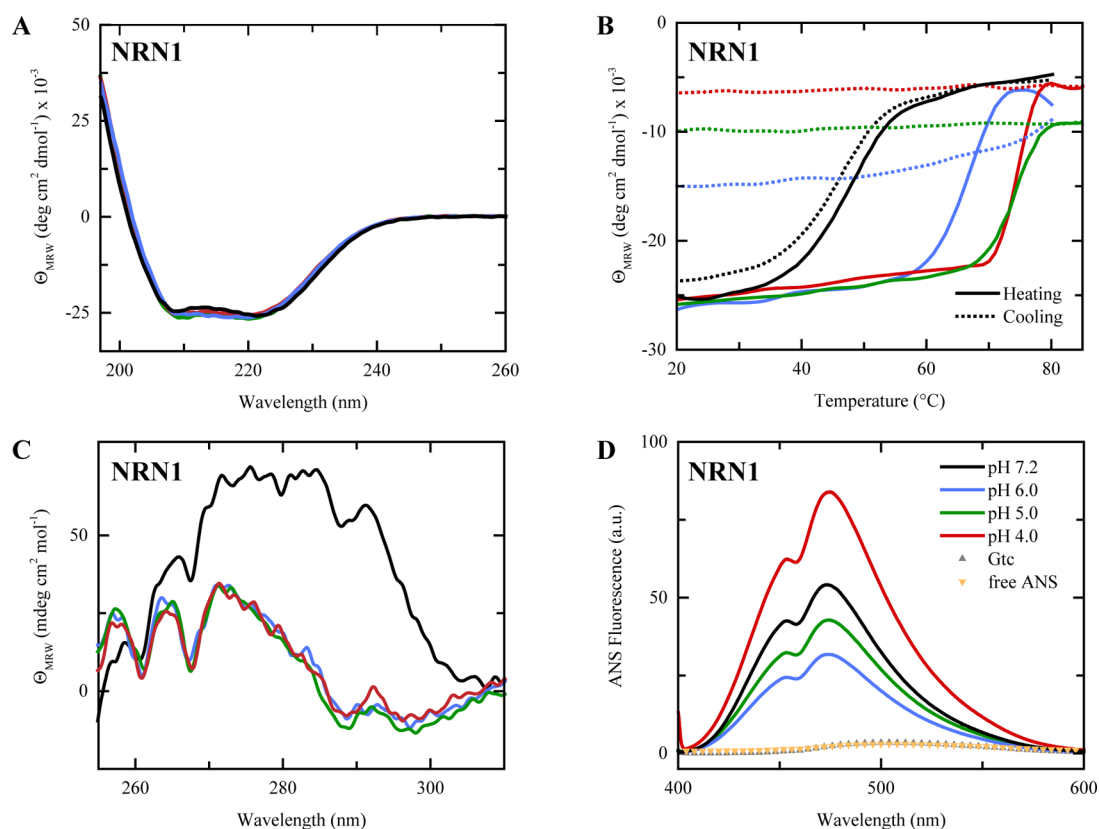
**CD Spectroscopy.** Circular dichroism experiments were performed on a Jasco J-715 spectropolarimeter using a quartz cuvette with a 1 mm path length. The scan speed was set to 50 nm min<sup>-1</sup>, response time to 1 s, acquisition interval to 0.1 nm, and bandwidth to 1 nm. The spectra shown were calculated as an average of three measurements and were subsequently smoothed by applying a Savitzky–Golay filter.<sup>32</sup> The secondary structure content was estimated by using JASCO Spectra Manager II and matching the spectra to Yang's references.<sup>33</sup> Far-UV CD spectra and thermal transitions were taken at a protein concentration of 0.2 mg mL<sup>-1</sup>, corresponding to 14.2  $\mu$ M NRN1 and 16.7  $\mu$ M NRC1, respectively. Temperature scans were monitored at a wavelength of 222 nm by recording at 1 °C steps and applying a heating rate of 1 °C min<sup>-1</sup>. For near-UV spectra, a path length of 0.5 cm and a concentration of 142.0  $\mu$ M NRN1 and 167.3  $\mu$ M NRC1 were applied.

**Fluorescence Spectroscopy.** 8-Anilino-naphthalene-1-sulfonic acid (ANS) binding to the spidroin domains was measured as a function of pH. After incubation of 250  $\mu$ M ANS and 1–40  $\mu$ M protein for 10 min at distinct pH values, fluorescence was monitored using an excitation wavelength of 390 nm, and emission spectra were taken between 400 and 600 nm employing a Jasco FP-6500 spectrofluorometer. Thioflavine T (ThT) binding was measured at 5 and 40  $\mu$ M protein and an excitation wavelength of 440 nm, and emission spectra were recorded between 450 and 600 nm. ThT and ANS fluorescence spectra are an average of three consecutive scans. For time-dependent measurements, ThT fluorescence was detected at an emission wavelength of 482 nm in intervals of 10 min.

**Surface Plasmon Resonance.** Surface plasmon resonance allows detection of dynamic protein–protein interactions in the nanomolar range.<sup>34,35</sup> A NTA sensor chip (GE Healthcare) was used in a Biocore X system (GE Healthcare), and a flow rate of 10  $\mu$ L min<sup>-1</sup> was applied. The chip surface was washed with 30  $\mu$ L of a 350 mM EDTA solution (pH 8.3) and activated by injection of 10  $\mu$ L of 0.5 mM NiSO<sub>4</sub>. Filtered HEPES (10 mM HEPES, pH 7.2, 150 mM NaCl) or NaP (10 mM NaP, pH 6.0) containing 0.1 mM EDTA and 0.005% Tween was used as buffer. To analyze the interaction between NRN1 and NRC1, 10  $\mu$ L of 0.1  $\mu$ M SUMO-tagged protein (bait) was applied to the second flow cell (FC2). Subsequently, FC1 was prepared to quantify unspecific interactions with control proteins (esterase-2: 34.3 kDa, pI = 5.1; SUMO: 13.4 kDa, pI = 6.0). As soon as stable signals were obtained for both FCs, the second protein (prey) was injected at molar ratios between 1 and 10 (10  $\mu$ L of 0.1–1  $\mu$ M).

**Turbidity Measurements.** Aggregation of NRC1 was analyzed as a function of pH by monitoring the UV absorption. The protein was diluted to the desired concentration by addition of 20 mM NaP (pH 7.2, pH 6.0), 20 mM NaOAc (pH 5.0, pH 4.0, pH 3.0), or 200 mM potassium phosphate buffer (pH adjusted with phosphoric acid). The





**Figure 2.** Analysis of pH-dependent structural changes of NRN1. The color coding for the different pH values is shown in panel D. (A) Far-UV CD spectra of 14.2  $\mu\text{M}$  NRN1 are indistinguishable between pH 7.2 and 4.0. (B) Thermal denaturation (solid line) and renaturation (dotted) of 14.2  $\mu\text{M}$  NRN1 monitored using CD ellipticity at 222 nm at pH values between 7.2 and 4.0. (C) Residual molar ellipticity of 142.5  $\mu\text{M}$  NRN1 determined in the near-UV region at pH values between 7.2 and 4.0. (D) Fluorescence emission spectra of 250  $\mu\text{M}$  ANS measured in the presence of 40  $\mu\text{M}$  NRN1 at pH values between 7.2 and 4.0. Emission spectra of ANS in the absence of protein (free ANS) and in the presence of denatured protein (6 M guanidinium thiocyanate, Gtc) served as negative controls.

absorbance was recorded at a wavelength of 330 nm in intervals of 10 min using a Varian Cary 50 UV–vis spectrophotometer (Agilent Technologies). The average of three measurements is shown together with maximal and minimal values.

**Congo Red Binding.** Congo red binding was performed as described previously.<sup>36</sup> Congo red was dissolved in buffer containing 150 mM NaCl and 5 mM  $\text{KH}_2\text{PO}_4$  (pH 7.4) and filtered three times using a 0.20  $\mu\text{m}$  filter (Sarstedt). Congo red concentration was determined by measuring the absorbance at 498 nm using an extinction coefficient of  $\epsilon = 3.4 \times 10^4 \text{ M}^{-1} \text{ cm}^{-1}$ . Proteins were incubated for 24 h at distinct pH values, the resulting aggregates were diluted with 20 mM NaP (pH 7.2) to 3  $\mu\text{M}$ , and Congo red was added to a final concentration of 4  $\mu\text{M}$ . After the mixture was incubated for 30 min, the absorbance was recorded between 400 and 600 nm, and the background absorbance of the protein was subtracted.

**SDS Stability Analysis.** SDS-PAGE sample buffer containing 1% (w/v) SDS was added to 2 mg  $\text{mL}^{-1}$  preaggregated protein. Protein samples were heated for 10 min at 95  $^\circ\text{C}$ , and 2.5  $\mu\text{g}$  of the protein samples was loaded onto a precast 4–20% acrylamide gradient gel (SERVAGel TG PRiME). After SDS-PAGE, the gel was silver stained.

**Scanning Electron Microscopy (SEM).** NRC1 aggregates were mounted onto a silica wafer by pipetting droplets of protein solution and drying overnight. Mounted samples were then sputtered with 1.3 nm platinum. Scanning electron microscopy (SEM) was performed using a Leo 1530 Gemini (Zeiss) equipped with InLens SE and SE2 detectors, applying an accelerating voltage of 3 kV.

**Transmission Electron Microscopy (TEM).** For transmission electron microscopy (TEM) experiments, 10  $\mu\text{L}$  of a preaggregated NRC1 suspension (418.4  $\mu\text{M}$ ) was spotted on supports (Pioloform 100-mesh copper grids and Formvar 200-mesh copper grids (Plano GmbH)) and incubated for 10 min. Samples were washed with 10  $\mu\text{L}$

of ddH<sub>2</sub>O (Millipore) and negatively stained upon incubation with 5  $\mu\text{L}$  of 2% uranyl acetate solution for 1 min. After two additional washing steps with 10  $\mu\text{L}$  of ddH<sub>2</sub>O, the samples were dried for at least 24 h at RT. A JEM-2100 transmission electron microscope (JEOL) was applied at 80 kV, and images were taken using a 4000  $\times$  4000 charge-coupled device camera (UltraScan 4000; Gatan) and Gatan Digital Micrograph software (version 1.83.842).

**Atomic Force Microscopy (AFM).** Atomic force microscopy (AFM) samples were prepared by pipetting drops of 10  $\mu\text{L}$  of preaggregated NRC1 (83.7  $\mu\text{M}$ ) onto freshly cleaved mica plates (L 10 mm, V1 grade, Plano GmbH). After incubation for 10 min, the samples were washed four times with 20  $\mu\text{L}$  of ddH<sub>2</sub>O (Millipore) and dried before imaging for at least 24 h at RT. For AFM scanning a Dimension 3100 equipped with a NanoScope V controller (Veeco Instruments Inc.) was employed, operating in TappingMode using  $\text{Si}_3\text{N}_4$  cantilevers (OMCL-AC160TS, Olympus, spring constant of 42 N/m, resonance frequency of 300 kHz, tip radius less than 7 nm). All images were acquired around the amplitude set point  $\sim 0.7$ – $0.9$ , and scans were evaluated using NanoScope Analysis software, version 7.20 (Bruker).

**Optical Microscopy and Photography.** The pH of a 5.0 mg  $\text{mL}^{-1}$  (AQ)<sub>24</sub>, (AQ)<sub>12</sub>-NR3, or (AQ)<sub>12</sub>-NRC1 solution was adjusted to either pH 7.2, 6.0, 5.0, or 4.0, and the samples were incubated for 18 h at RT while rotating at 25 rpm. The aggregates were transferred onto glass slides for bright-field microscopy (DMI3000B, Leica) using a 40 $\times$  objective lens. For photography, the samples were placed in glass tubes, and pictures were taken with a Nikon D5100.

## RESULTS

**Structural Analysis of Individual Terminal Spidroin Domains.** The structure of the nonrepetitive NRN1 and NRC1 spidroin domains of *L. hesperus* MaSp1 and spidroin assembly were both characterized under various pH conditions, including nonphysiological acidic ones. Far-UV CD measurements indicated a stable  $\alpha$ -helical secondary structure of NRN1 between pH 7.2 and pH 4.0 (Figure 2A). Importantly, pH 4.0 is far below physiological conditions, as the lowest pH measured in any spider's spinning duct thus far has been 5.7.<sup>25</sup> The thermal stability significantly increased upon acidification, from 47 to 74 °C (Figure 2B and Table 1). Acidification below pH

**Table 1.** pH-Dependent Melting Temperatures ( $T_m$ ) of NRC1, NRN1, and 1:1 Mixtures<sup>a</sup>

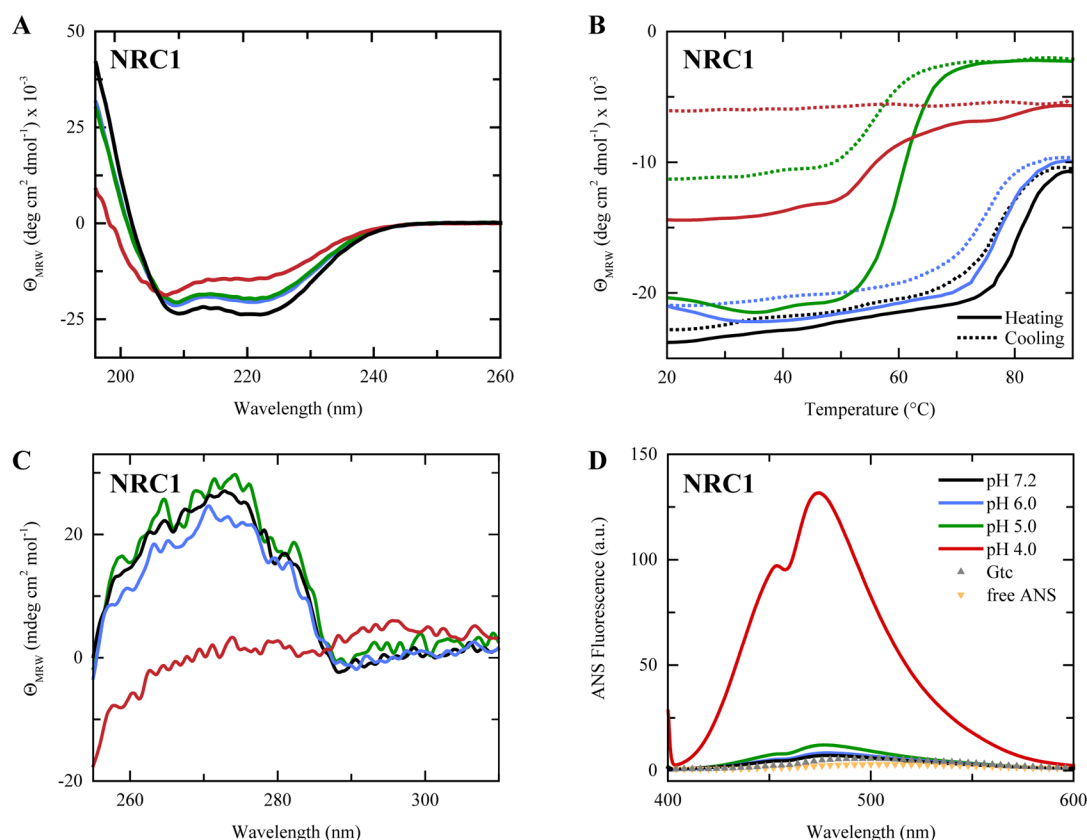
$T_m$	NRN1	NRC1	NRN1 + NRC1
pH 7.2	47 °C	82 °C	47 °C, 82 °C
pH 6.0	66 °C*	77 °C	59 °C*, 78 °C
pH 5.0	73 °C	60 °C	70 °C
pH 4.0	74 °C	56 °C	56 °C, 72 °C

<sup>a</sup>Thermal denaturation of 0.2 mg mL<sup>-1</sup> protein was analyzed using CD ellipticity at 222 nm. Theoretical isoelectric points of 4.1 and 4.5 were predicted for NRN1 and NRC1, respectively.  $T_m$  of the individual NRN1 and the mixture at pH 6.0 was found to be significantly different (marked by \*). At pH 5.0, the overlapping melting curves of the mixture merged into one transition (shown in italics).

4.0 had no further influence on the observed properties (data not shown). Strikingly, refolding after thermal unfolding was completely suppressed at pH values below 6.0 (Figure 2B). Near-UV CD spectra indicated that the environment of the aromatic amino acid side chains, i.e., the tertiary structure, was significantly changed between pH 7.2 and 6.0 and stable below that pH (Figure 2C).

In general, ANS can be used to analyze exposition of hydrophobic patches, typically indicating partial unfolding of a protein.<sup>37</sup> Here, NRN1 showed the lowest affinity toward ANS at pH 6.0, resulting from the reorganization of the tertiary structure (Figure 2D) and dimer formation (Figure S1).<sup>26</sup> Further acidification increased the ANS emission, indicating the exposure of hydrophobic patches at the protein surface (Figure 2D). To test whether the exposure of hydrophobic patches induced aggregation, the properties were tested in a concentration-dependent manner. The ANS fluorescence linearly correlated with protein concentration between pH 7.2 and 4.0 (Figure S2), suggesting the lack of aggregation. This was independently confirmed by turbidity measurements (Figure S2).

In contrast to NRN1, NRC1's content of  $\alpha$ -helices significantly decreased from 56 to 17% upon acidification, as seen by far-UV CD measurements (Figure 3A, Table S1). In accordance, the thermal stability (Figure 3B and Table 1) of NRC1 was decreased. Similar to NRN1, the thermal renaturation of the solution structure of NRC1 was completely

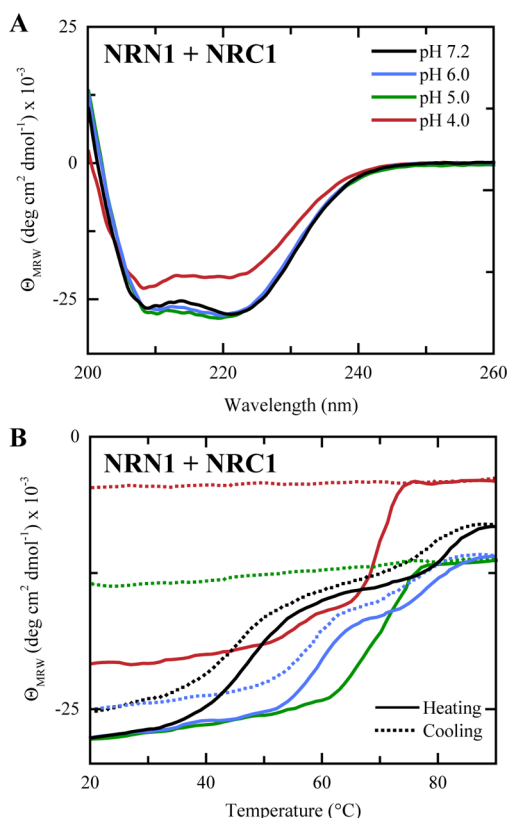


**Figure 3.** Structural analysis of NRC1. The color coding for the different pH values is shown in panel D. (A) Far-UV CD spectra of 16.7  $\mu$ M NRC1, indicating the loss of helical content upon acidification. (B) Thermal denaturation (solid line) and renaturation (dotted) of 16.7  $\mu$ M NRC1 monitored using CD ellipticity at 222 nm at pH values between 7.2 and 4.0. (C) Residual molar ellipticity of 167.3  $\mu$ M NRC1 in the near-UV region at pH values between 7.2 and 4.0. (D) Fluorescence emission spectra of 250  $\mu$ M ANS recorded in the presence of 40  $\mu$ M NRC1 at pH values between 7.2 and 4.0.



suppressed below pH 5.0 (Figure 3B). The tertiary structure of NRC1 was maintained until pH 5.0, but below that pH the aromatic amino acid residues were no longer embedded in a stable structural environment (Figure 3C). This was confirmed by a higher affinity toward ANS at pH 4.0 (Figure 3D). Furthermore, acidic conditions triggered NRC1 aggregation, as indicated by a slightly logarithmic correlation between ANS fluorescence and protein concentration (Figure S2). Turbidity measurements confirmed that NRC1 has a strong tendency to form aggregates at nonphysiological pH 4.0 (Figure S2). Taken together, acidification triggered slightly opposite effects in both domains, converting NRN1 into a stabilized dimer and NRC1 into a destabilized conformation.

**Structural Analysis of NRN1–NRC1 Mixtures.** So far, in all previous studies the terminal spidroin domains were investigated individually.<sup>6,7,22,25–27,30</sup> To test putative intermolecular interactions between both domains at neutral and acidic pH, here the conformational and thermal stabilities of NRN1–NRC1 mixtures were analyzed using CD spectroscopy (Figure 4). Consistent with spectra of both individual five-helix



**Figure 4.** Structural analysis of 1:1 mixtures of NRN1 and NRC1. The color coding for the different pH values is shown in panel A. (A) Far-UV CD spectra of 0.1 mg mL<sup>−1</sup> NRN1 and 0.1 mg mL<sup>−1</sup> NRC1 at pH values between 7.2 and 4.0. (B) Thermal denaturation (solid line) and renaturation (dotted) of NRN1–NRC1 mixtures measured using CD ellipticity at 222 nm at pH values between 7.2 and 4.0.

bundles (Figures 2A and 3A), the mixtures exhibited a high  $\alpha$ -helical content at neutral pH (Figure 4A). At pH 4.0, the  $\alpha$ -helical signal of the mixtures decreased from 52 to 26% (Table S2), indicating the destabilization of NRC1 in the mixture at low pH (Figure 4). The  $T_m$  values of NRN1 and NRC1 were unchanged after mixing (Table 1), with the exception of a slight

thermal destabilization at pH 6.0 ( $\Delta T_m = -7^\circ\text{C}$ ), which could be the result of possible protein–protein interactions.

**Interaction Analysis between NRN1 and NRC1.** To analyze putative domain–domain interactions in more detail, independent experiments using a pull-down assay, native PAGE, SPR, and multiangle light scattering (MALS) analysis were performed. The pull-down assay indicated that NRC1 did not specifically interact with SUMO–NRN1 under solvent conditions representing storage conditions in the spinning sac (200 mM NaCl, pH 7.2) or under assembly-triggering conditions as found in the spinning duct (200 mM K<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub>, pH 6.0) (Figure S3). Since fusion of SUMO with NRN1 and binding to the beads might sterically hinder interactions with the partner protein, the interplay between nonimmobilized NRN1 and NRC1 was independently examined by native PAGE (Figure S3) and SEC-MALS (Figure S1), but again the terminal spidroin domains apparently did not interact under solvent conditions found during spidroin storage in the sac or assembly in the duct.

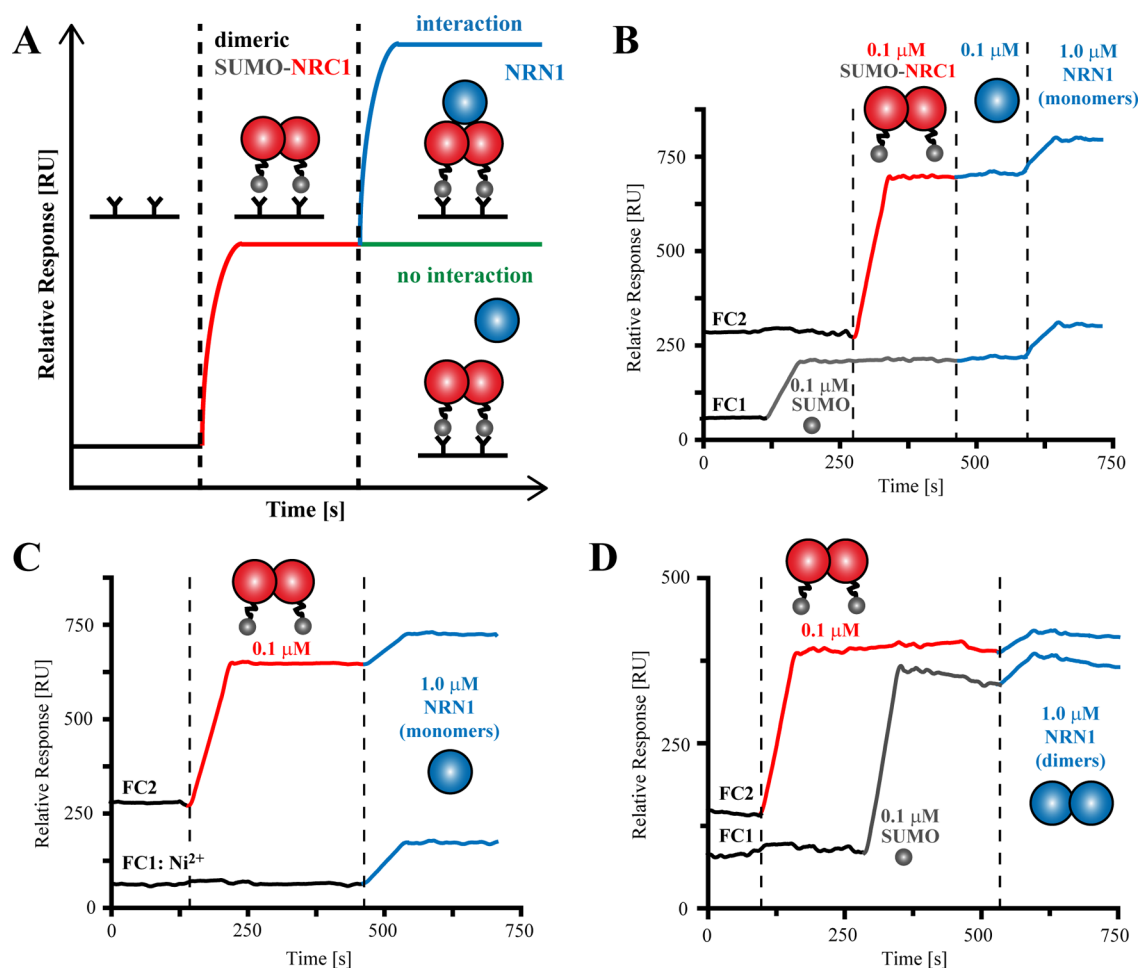
Since pull-down and native PAGE analyses are providing relatively low resolution and are not applicable to detect dynamic short-time interactions, SPR was used to exploit putative interactions with high sensitivity and to monitor dynamic interactions. A NTA sensor chip was activated with Ni<sup>2+</sup> ions before immobilizing SUMO-tagged protein and injecting the analyte (Figure 5A). Two independent experimental setups were used in which the immobilized proteins were switched, testing the interplay between immobilized SUMO–NRC1 and soluble NRN1 (Figure 5B,C) and between immobilized SUMO–NRN1 and soluble NRC1 (Figure S4). The other flow cell (FC1) of the sensor chip was used as reference to monitor nonspecific interactions of analytes toward immobilized protein controls (plain SUMO and esterase-2) or Ni<sup>2+</sup> ions.

Immobilization of SUMO–NRC1 to the surface of FC2 was visualized by a binding response of approximately 500 RU (Figure 5B). The response remained identical when NRN1 was injected additionally, indicating the lack of specific interaction between both proteins in this assay. Increasing the injected NRN1 concentration to a 10-fold molar excess increased proportionally the nonspecific interaction with all reactants, SUMO–NRC1, SUMO, or Ni<sup>2+</sup> (Figure 5B,C). Dimerization of NRN1 at pH 6.0 also did not enhance the affinity toward SUMO–NRC1 (Figure 5D).

The control esterase-2 resembles dimeric NRN1 in terms of MW (34.3 kDa) and pI (5.1).<sup>38,39</sup> Therefore, it was used as a reference analyte, verifying the nonspecific binding to SUMO–NRC1 or a Ni<sup>2+</sup>-covered surface (Figure S4). To validate the SPR setup, NRC1 was cleaved from immobilized SUMO–NRC1 by injecting SUMO protease (Figure S4). The initial baseline was reached after cleaving off NRC1, confirming that the technique provides the sensitivity to detect changes in the given MW range.

#### Apparent Differences between Amyloids and NRC1 Aggregates.

Since a destabilization of NRC1 at acidic pH values was detected, the implication of such an event on spidroin assembly was evaluated. Previously, below pH 5.5, carboxy-terminal domains of *N. clavipes* MaSp1 and *A. ventricosus* MiSp have been suggested to convert into  $\beta$ -sheet-rich nuclei, triggering assembly into amyloid-like conformations.<sup>25,30</sup> Therefore, NRC1 of *L. hesperus* MaSp1 was analyzed with regard to  $\beta$ -sheet formation at low pH by applying dyes typically used for amyloid detection, such as Congo red and



**Figure 5.** SPR analysis of the interaction between dimeric SUMO–NRC1 and monomeric/dimeric NRN1. (A) Schematic chromatogram of a SPR measurement. (B) Affinity of 0.1 and 1.0  $\mu\text{M}$  NRN1 monomers to 0.1  $\mu\text{M}$  SUMO–NRC1 dimers (FC2) and to 0.1  $\mu\text{M}$  SUMO (control, FC1). (C) Binding of 1.0  $\mu\text{M}$  NRN1 to 0.1  $\mu\text{M}$  SUMO–NRC1 (FC2) is indistinguishable to that of unspecific binding to a  $\text{Ni}^{2+}$  loaded flow cell (FC1). (D) Dimerization of NRN1 at pH 6.0 did not enhance the affinity toward SUMO–NRC1 (FC2) and SUMO (FC1).

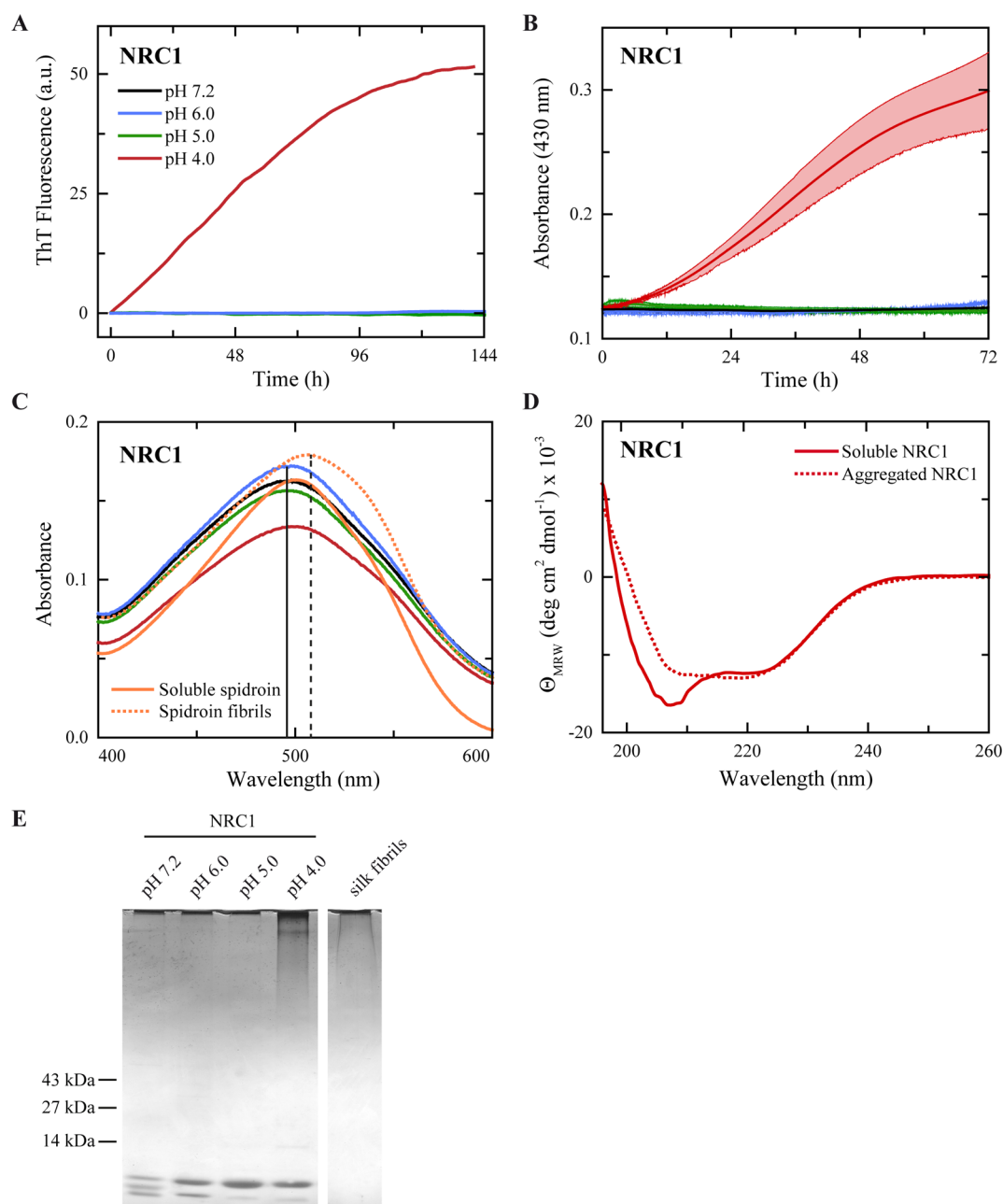
ThT, and testing the structural resistance against SDS. NRC1 aggregation, induced by acidification, was time-dependently recorded by turbidity measurements. The resulting structures were microscopically investigated by SEM, TEM, and AFM.

Previously, the carboxy-terminal domain of *A. ventricosus* MiSp1 was shown to interact with ThT,<sup>25</sup> a fluorescent dye that enhances its fluorescence emission upon binding to amyloid-like structures.<sup>40,41</sup> At neutral and slightly acidic pH ( $\geq 5.0$ ), no ThT binding to NRC1 of *L. hesperus* could be detected (Figure 6A). However, upon acidification beyond the pI of NRC1 (4.5) at pH 4.0 the intensity of ThT fluorescence slowly increased over 6 days.

Time-dependent turbidity measurements confirmed these findings (Figure 6B). At pH 4.0, the absorbance increased over 3 days, indicating the formation of aggregates. Increasing the temperature or protein concentration significantly supported this type of aggregation (Figure S5). The resulting NRC1 aggregates were further analyzed for Congo red binding, which is used for detecting  $\beta$ -sheets in amyloid fibrils by red shifting the Congo red absorbance maximum upon binding.<sup>36,42</sup> Previously, it has been shown that Congo red bound to fibrils made of recombinant spidroins comprising only sequences of the core domain (eADF4(C16)) derived from *Araneus diadematus* fibroin 4, which was confirmed here (Figure 6C).<sup>31</sup> Since the absorbance of protein–Congo red mixtures

was slightly affected by the pH value (Figure S6), the preoligomerized NRC1 samples were neutralized before the measurement. NRC1 aggregates formed at pH 4.0 showed a slight red shift of Congo red absorbance by 4.8 nm (Figure 6C), indicating that the  $\beta$ -sheet content increased upon acidification, as confirmed by far-UV CD spectra (Figure 6D). Since the Congo red absorbance maximum did not reach a wavelength of 506 nm, as seen for eADF4(C16) fibrils, the NRC1 aggregates definitely exhibited a lower content and/or less orientation of  $\beta$ -sheets than in spidroin fibrils. SDS analysis, in which normal protein aggregates dissolve but  $\beta$ -sheet-rich amyloids are stable,<sup>43</sup> indicated that monomeric and dimeric NRC1 were solubilized. At pH 4.0, the amount of dissolved protein slightly decreased, and oligomers with larger MW formed (Figure 6E). However, since highly ordered fibrils consisting of eADF4(C16) were completely insoluble in SDS (Figure 6E), the acidic-triggered aggregates of NRC1 can be considered to be significantly less rigid than amyloid-like fibers; in other words, they do not show amyloid-like features.

**Effect of NRC1 Aggregates on Soluble NRN1.** One further feature of amyloid-like structures is their ability to catalyze the formation of new amyloid structures within soluble counterparts (a phenomenon called seeding and cross-seeding). Although the folding of the terminal domains was not affected in the presence of each other and no heterogeneous protein



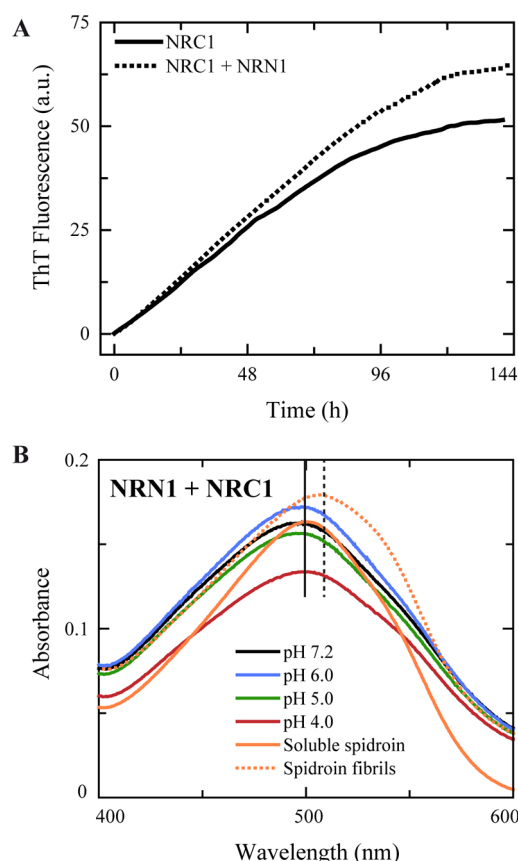
**Figure 6.** Analysis of pH-dependent NRC1 aggregation. The color coding for the different pH values is shown in panel A. (A) Fluorescence of  $5\ \mu\text{M}$  ThT monitored at a wavelength of  $482\ \text{nm}$  in the presence of  $40\ \mu\text{M}$  NRC1 at pH values between 7.2 and 4.0 over time. (B) Turbidity of  $83.7\ \mu\text{M}$  NRC1 recorded at a wavelength of  $330\ \text{nm}$  in a time-dependent manner ( $n = 3$ ). Here, the effect of pH was tested in salt-free buffers. (C) After incubation of  $3\ \mu\text{M}$  NRC1 for 1 week at distinct pH values, the binding of  $4\ \mu\text{M}$  Congo red was recorded. Congo red binding of amyloid-like spidroin fibrils (eADF4(C16))<sup>31</sup> is shown as a control, providing a significant red shift in comparison to that of NRC1 aggregates. (D) Far-UV CD spectra of soluble NRC1 recorded at a concentration of  $16.7\ \mu\text{M}$  at pH 4.0. Samples were subsequently incubated for 1 week at RT, and spectra were taken of the resulting aggregates. (E) SDS solubility of aggregated NRC1, as observed by PAGE. Prior to addition of 1% SDS and boiling for 10 min, the samples were incubated at distinct pH values for 6 days. Spidroin fibrils of eADF4(C16) are analyzed as a control.

complexes were formed, it is still possible that NRC1 amyloids could act as nucleation sites/seeds for NRN1 aggregation. Therefore, mixtures of both proteins were analyzed for ThT and Congo red binding (Figure 7). In contrast to NRC1, NRN1 had no affinity toward ThT in the tested pH range (Figure S7). Addition of NRN1 to NRC1 did not significantly accelerate  $\beta$ -sheet formation, but it slightly increased the overall ThT binding (Figure 7A). Congo red binding to NRC1 was not altered in the presence of NRN1 (Figure 7B). Hence, the  $\beta$ -sheet formation of NRC1 at low pH was not significantly

affected by NRN1; otherwise, it barely seeded the aggregation of NRN1.

#### Macroscopic Characterization of NRC1 Aggregates.

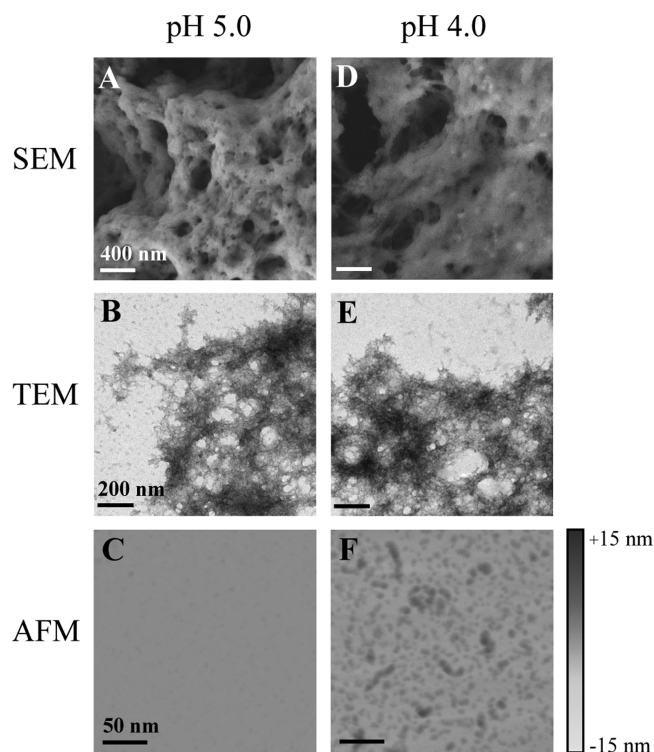
Finally, the macroscopic changes of NRC1 aggregates at different pH conditions were analyzed using SEM, TEM, and AFM. At pH 5.0, nanometer-sized globular aggregates accumulated into a porous and heterogeneous morphology of up to a few hundred nanometers in size, as seen in SEM and TEM images (Figure 8A,B). The aggregates were almost completely resolubilized upon 5-fold dilution in buffer (pH



**Figure 7.** Effect of the presence of NRN1 on NRC1 aggregation. (A) Fluorescence of 5  $\mu\text{M}$  ThT monitored at a wavelength of 482 nm over time in the presence of 40  $\mu\text{M}$  NRC1 and in the absence or presence of 40  $\mu\text{M}$  NRN1. (B) Absorbance of 4  $\mu\text{M}$  Congo red recorded in the presence of 3  $\mu\text{M}$  NRC1 and 3  $\mu\text{M}$  NRN1 after incubation for 1 week at distinct pH values. Fibril formation of recombinant spidroins (eADF4(C16)) illustrates the typical spectroscopic change of Congo red in the presence of amyloid-like structures.

5.0), as observed by AFM (Figure 8C). At pH 4.0, NRC1 formed globular aggregates with a height and width of approximately 10 nm (Figure 8F; surface roughness arithmetic average  $R_a = 1.10$  nm,  $R_{\text{max}} = 12.3$  nm). These spheres further agglomerated to form a disordered network of spherical and linear aggregates (Figure 8D,E). Although NRC1 exhibited some properties of amyloids such as an increased  $\beta$ -sheet content (Figure 6), the protein did not assemble into clearly defined fibrils, as is typically seen for amyloidogenic proteins.

**Distinction between Aggregation of NRC1 and Spidroin Assembly.** To correlate the aggregation of NRC1 with spidroin self-assembly, two homologue carboxy-terminal domains (NRC1 of *L. hesperus* MaSp1 and NR3 of *A. diadematus* MaSp1) were fused with a previously established engineered core domain (AQ)<sub>12</sub> based on the consensus sequence of *A. diadematus* fibroin 3 (ADF3).<sup>31</sup> (AQ)<sub>24</sub> was used as control, existing as a monomer with a MW comparable to that of dimers of (AQ)<sub>12</sub>-NRC1 or (AQ)<sub>12</sub>-NR3 (Figure S8 and Table S3). Exposure of (AQ)<sub>12</sub>-NRC1 to shear stress led to well-defined fibrous aggregates at pH values between 7.2 and 5.0 (Figure 9A). The micrometer to millimeter sized aggregates were composed of aligned bundles of fibrils, as observed by bright-field microscopy (Figure 9B).<sup>7</sup> In contrast, ill-defined aggregates with significantly smaller dimensions were formed at pH 4.0 (Figure 9A,C). The “loss-of-function” at low pH was



**Figure 8.** Microscopic images of pH-triggered NRC1 aggregates. The pH of a 418.4  $\mu\text{M}$  NRC1 solution was adjusted to pH 5.0 and 4.0, respectively, and the samples were incubated for 24 h prior to SEM and TEM. For AFM, the samples were diluted to 83.7  $\mu\text{M}$  with the corresponding buffer. The bar in the AFM images represents height from  $-15$  nm (bright) to  $+15$  nm (dark).

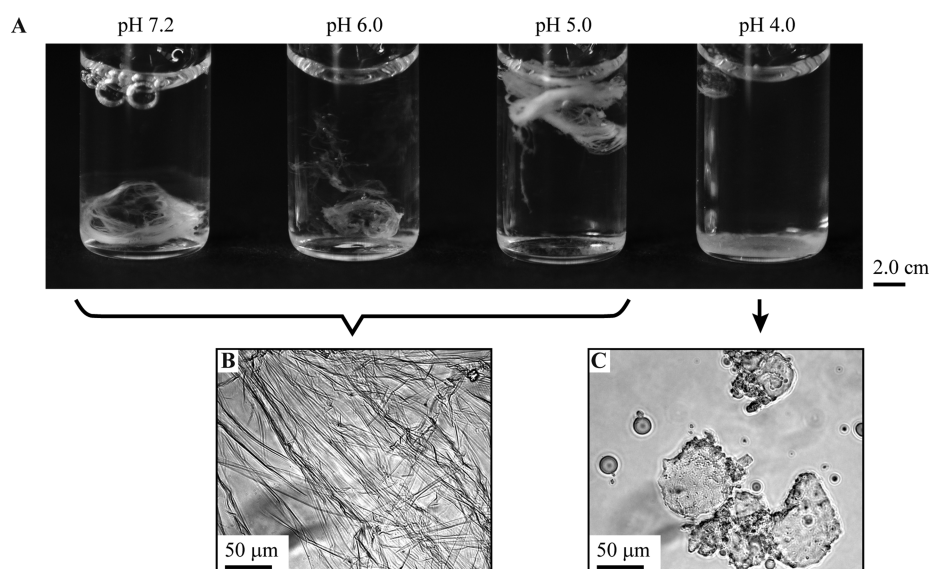
also observed for (AQ)<sub>12</sub>-NR3 (Figure S9), suggesting that this behavior might be not a unique feature of NRC1 of *L. hesperus*. The findings confirmed on the one hand that natively folded NRC aligns the spidroins into fibrous structures in response to shear stress during self-assembly, but on the other hand this ability is lost at low, nonphysiological pH.

In summary, the native structure of the carboxy-terminal domain of *L. hesperus* MaSp1 is substantially destabilized under acidic conditions, which favors the formation of  $\beta$ -sheet-rich aggregates. The properties of NRC1 were not affected by natively folded NRN1, but the extended conformation of NRC1 at low pH triggered recombinant spidroins to form unstructured aggregates with low rigidity rather than aligned assemblies as seen at physiologically relevant pH values.

## DISCUSSION

Prior to spinning, the highly concentrated silk protein (spidroin) solution is stored in the reservoir of a spider's gland. The amphiphilic spidroins assemble into micelle-like structures to prevent nonspecific aggregation.<sup>24,44</sup> Water solubility is promoted by the spidroin's hydrophilic non-repetitive NRN and NRC domains, which are both exposed at the surface of the micelles. Here, it was shown that the folding of neither NRN1 nor NRC1 of *L. hesperus* major ampullate spidroin 1 (MaSp1) was affected by the opposing domain (Figure 4), and they apparently do not interact (Figure 5). Nonetheless, the domains respond to progressive changes of their environment along the spinning duct, triggering the transition of soluble spidroins into solid fibers. Shear forces in combination with acidification induce the dimeric NRC to





**Figure 9.** Assembly of recombinant spidroins. (A) The pH of a 85.72  $\mu\text{M}$  (AQ)<sub>12</sub>-NRC1 solution was adjusted to 7.2, 6.0, 5.0, and 4.0. Samples were rotated overnight, and pictures of the resulting aggregates were taken. (B, C) Bright-field microscopy images of (AQ)<sub>12</sub>-NRC1 aggregated at pH 7.2 (B) and 4.0 (C).

slightly change its conformation.<sup>7,24,25,30</sup> In the case of NRN1, acidification induces the formation of antiparallel homodimers.<sup>6,26</sup>

The most stable structure of NRN1 was observed at pH 6.0, and acidification did not further affect the folding of this domain (Figure 2). Interestingly, the structure of NRN1 was slightly destabilized in the presence of NRC1, raising the question of a putative domain–domain interaction. However, neither stable nor dynamic short-time interactions could be detected between NRN1 and NRC1 in pull-down assays (Figure S3), native PAGE (Figure S3), SPR (Figure 5), and MALS (Figure S1) under solvent conditions similar to those found during spidroin storage in the sac or during assembly in the duct. Since the N-terminal domain switches its structure in response to the ionic environment and the C-terminal domain reacts in response to shear stress, this behavior could represent a double-safety mechanism in response to chemical and mechanical triggers to prevent premature aggregation and to enable the tight control of spidroin assembly within the spinning duct.

The tightly folded five-helix bundle of NRC1 is stabilized at physiological pH by interactions of the side chains including intramolecular salt bridges between acidic residues (D76, D104, E84) and a single arginine (R35). Acidification-induced protonation of the acidic side chain residues at nonphysiological pH 4.0 significantly destabilizes the native fold (Figure 3). Similarly, it was previously reported that the disruption of salt bridges at low pH induced the loss of structure of the NR3 domain of *A. diadematus* fibroin 3.<sup>7</sup> The structural destabilization caused NRC1 to form aggregates which bound ThT and Congo red to some extent (Figure 6). However, the kinetics of aggregation were slow (several days), making it unlikely for this behavior to be relevant for silk fiber formation, which occurs within milliseconds.<sup>45</sup> NRC1 aggregates further exhibited a low rigidity with clear differences from highly rigid  $\beta$ -sheet-rich silk fibers.<sup>12</sup> SEM micrographs as well as TEM and AFM images showed a predominantly heterogeneous and unstructured morphology of globular NRC1 aggregates (Figure 8), clearly different from those of previously analyzed homologue proteins

that form more elongated and nonbranched fibrils.<sup>25,30</sup> These results suggest that the slightly diverse behavior of NRC from different species at low pH indicates a rather species-specific feature, which is unlikely to be part of a general fiber assembly mechanism. Natively folded NRC1 controls the self-assembly of recombinant spidroins into fibrous structures, but this function is lost in the unfolded structure at low pH. Even so, NRC1 forms  $\beta$ -sheet-rich aggregates at very low (non-physiological) pH, this domain feature does not act as a seed for spidroin assembly into ordered fibers. The aggregation of NRC1 can be clearly distinguished from the highly dynamic, controlled, and reversible self-assembly of spidroins into well-defined fibers as well as from amyloid formation.

## CONCLUSIONS

It could be shown that the terminal spidroin domains react individually to external stimuli and undergo opposing conformational changes upon acidification. Low pH stabilized the structure of NRN1, but it triggered NRC1 to unfold and aggregate. Even so, NRC1 aggregates formed at nonphysiological pH exhibited some characteristics of amyloids such as increased  $\beta$ -sheet content, and the nonspecific aggregation could be clearly separated from amyloid and spidroin self-assembly into aligned fibers. The results imply that each terminal domain plays a significant, independent role in the prevention of protein aggregation within the gland, as well as in fiber assembly of spider dragline silk. This work contributes new details about the molecular features of spider silk proteins/domains, which can help to establish biomimetic spinning technologies. Under physiological conditions, the soluble and natively folded terminal domains act as individual molecular switches controlling the self-assembly of spidroins into highly ordered structures; however, their natural function is lost at low, nonphysiological pH.

## ■ ASSOCIATED CONTENT

### ■ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biomac.6b01713.

Aggregation analysis of NRN1 and NRC1; SPR chromatograms; pH-, protein- concentration-, and temperature-dependent turbidity measurements of NRC1; pH-dependent absorbance of Congo red; fluorescence of Thioflavin T in the presence of NRN1; dimerization analysis of recombinant spidroins; bright-field microscopy images of recombinant spidroin aggregates; and secondary structure estimations for NRC1 and for mixtures of NRN1 and NRC1 as well as characteristics of recombinant spidroins (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: thomas.scheibel@bm.uni-bayreuth.de.

### ORCID

Thomas Scheibel: 0000-0002-0457-2423

### Author Contributions

Conceptualization, J.B. and T.S.; investigation, J.B.; writing (original draft), J.B.; writing (review and editing), T.S.; visualization, J.B.; supervision, T.S.

### Funding

Funding was obtained from DFG (SFB 840 TP A8).

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We would like to thank Dr. Martin Humenik for experimental help and AFM measurements, Christian Borkner for performing SEM, Tamara Aigner as well as Kristin Schacht for help with TEM, and Corinna Drescher for performing native PAGE and SEC-MALS analysis. Furthermore, we would like to thank Dr. Martin Humenik and Elise Desimone for proof reading.

## ■ REFERENCES

- (1) Vollrath, F. Strength and structure of spiders' silks. *Rev. Mol. Biotechnol.* **2000**, *74* (2), 67–83.
- (2) Blackledge, T. A.; Hayashi, C. Y. Silken toolkits: biomechanics of silk fibers spun by the orb web spider *Argiope argentata* (Fabricius 1775). *J. Exp. Biol.* **2006**, *209* (13), 2452–2461.
- (3) Madsen, B.; Shao, Z. Z.; Vollrath, F. Variability in the mechanical properties of spider silks on three levels: interspecific, intraspecific and intraindividual. *Int. J. Biol. Macromol.* **1999**, *24* (2–3), 301–306.
- (4) Simmons, A. H.; Michal, C. A.; Jelinski, L. W. Molecular orientation and two-component nature of the crystalline fraction of spider dragline silk. *Science* **1996**, *271* (5245), 84–7.
- (5) Hijirida, D. H.; Do, K. G.; Michal, C.; Wong, S.; Zax, D.; Jelinski, L. W. <sup>13</sup>C NMR of *Nephila clavipes* major ampullate silk gland. *Biophys. J.* **1996**, *71* (6), 3442–3447.
- (6) Askarieh, G.; Hedhammar, M.; Nordling, K.; Saenz, A.; Casals, C.; Rising, A.; Johansson, J.; Knight, S. D. Self-assembly of spider silk proteins is controlled by a pH-sensitive relay. *Nature* **2010**, *465* (7295), 236–238.
- (7) Hagn, F.; Eisoldt, L.; Hardy, J. G.; Vendrely, C.; Coles, M.; Scheibel, T.; Kessler, H. A conserved spider silk domain acts as a molecular switch that controls fibre assembly. *Nature* **2010**, *465* (7295), 239–242.

- (8) Ayoub, N. A.; Garb, J. E.; Tinghitella, R. M.; Collin, M. A.; Hayashi, C. Y. Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. *PLoS One* **2007**, *2* (6), e514.
- (9) Hinman, M. B.; Lewis, R. V. Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber. *J. Biol. Chem.* **1992**, *267* (27), 19320–19324.
- (10) Guerette, P. A.; Ginzinger, D. G.; Weber, B. H.; Gosline, J. M. Silk properties determined by gland-specific expression of a spider fibroin gene family. *Science* **1996**, *272* (5258), 112–115.
- (11) Sanggaard, K. W.; Bechsgaard, J. S.; Fang, X. D.; Duan, J. J.; Dyrland, T. F.; Gupta, V.; Jiang, X. T.; Cheng, L.; Fan, D. D.; Feng, Y.; Han, L. J.; Huang, Z. Y.; Wu, Z. Z.; Liao, L.; Settepani, V.; Thogersen, I. B.; Vanthournout, B.; Wang, T.; Zhu, Y. B.; Funch, P.; Enghild, J. J.; Schauer, L.; Andersen, S. U.; Villesen, P.; Schierup, M. H.; Bilde, T.; Wang, J. Spider genomes provide insight into composition and evolution of venom and silk. *Nat. Commun.* **2014**, *5*, 3765.
- (12) Thiel, B. L.; Viney, C. Beta sheets and spider silk. *Science* **1996**, *273* (5281), 1480–1481.
- (13) Brooks, A. E.; Steinkraus, H. B.; Nelson, S. R.; Lewis, R. V. An investigation of the divergence of major ampullate silk fibers from *Nephila clavipes* and *Argiope aurantia*. *Biomacromolecules* **2005**, *6* (6), 3095–3099.
- (14) Liu, Y.; Spönnner, A.; Porter, D.; Vollrath, F. Proline and processing of spider silks. *Biomacromolecules* **2008**, *9* (1), 116–121.
- (15) Perez-Rigueiro, J.; Elices, M.; Plaza, G. R.; Guinea, G. V. Similarities and differences in the supramolecular organization of silkworm and spider silk. *Macromolecules* **2007**, *40* (15), 5360–5365.
- (16) Knight, D. P.; Vollrath, F. Liquid crystals and flow elongation in a spider's silk production line. *Proc. R. Soc. London, Ser. B* **1999**, *266* (1418), 519–523.
- (17) Bini, E.; Knight, D. P.; Kaplan, D. L. Mapping domain structures in silks from insects and spiders related to protein assembly. *J. Mol. Biol.* **2004**, *335* (1), 27–40.
- (18) Rising, A.; Hjalm, G.; Engstrom, W.; Johansson, J. N-terminal nonrepetitive domain common to dragline, flagelliform, and cylindri-form spider silk proteins. *Biomacromolecules* **2006**, *7* (11), 3120–4.
- (19) Garb, J. E.; Ayoub, N. A.; Hayashi, C. Y. Untangling spider silk evolution with spidroin terminal domains. *BMC Evol. Biol.* **2010**, *10*, 243.
- (20) Chen, G. F.; Liu, X. Q.; Zhang, Y. L.; Lin, S. Z.; Yang, Z. J.; Johansson, J.; Rising, A.; Meng, Q. Full-Length Minor Ampullate Spidroin Gene Sequence. *PLoS One* **2012**, *7* (12), e52293.
- (21) Otikovs, M.; Chen, G.; Nordling, K.; Landreh, M.; Meng, Q.; Jörnvall, H.; Kronqvist, N.; Rising, A.; Johansson, J.; Jaudzems, K. Diversified Structural Basis of a Conserved Molecular Mechanism for pH-Dependent Dimerization in Spider Silk N-Terminal Domains. *ChemBioChem* **2015**, *16* (12), 1720–1724.
- (22) Hagn, F.; Thamm, C.; Scheibel, T.; Kessler, H. pH-Dependent Dimerization and Salt-Dependent Stabilization of the N-terminal Domain of Spider Dragline Silk—Implications for Fiber Formation. *Angew. Chem., Int. Ed.* **2011**, *50* (1), 310–313.
- (23) Knight, D. P.; Vollrath, F. Changes in element composition along the spinning duct in a *Nephila* spider. *Naturwissenschaften* **2001**, *88* (4), 179–82.
- (24) Vollrath, F.; Knight, D. P. Liquid crystalline spinning of spider silk. *Nature* **2001**, *410* (6828), 541–548.
- (25) Andersson, M.; Chen, G. F.; Otikovs, M.; Landreh, M.; Nordling, K.; Kronqvist, N.; Westermark, P.; Jörnvall, H.; Knight, S.; Ridderstrale, Y.; Holm, L.; Meng, Q.; Jaudzems, K.; Chesler, M.; Johansson, J.; Rising, A. Carbonic Anhydrase Generates CO<sub>2</sub> and H<sup>+</sup> That Drive Spider Silk Formation Via Opposite Effects on the Terminal Domains. *PLoS Biol.* **2014**, *12* (8), e1001921.
- (26) Bauer, J.; Schaal, D.; Eisoldt, L.; Schweimer, K.; Schwarzwinger, S.; Scheibel, T. Acidic Residues Control the Dimerization of the N-terminal Domain of Black Widow Spiders' Major Ampullate Spidroin 1. *Sci. Rep.* **2016**, *6*, 34442.
- (27) Jaudzems, K.; Askarieh, G.; Landreh, M.; Nordling, K.; Hedhammar, M.; Jörnvall, H.; Rising, A.; Knight, S. D.; Johansson, J. pH-dependent dimerization of spider silk N-terminal domain

requires relocation of a wedged tryptophan side chain. *J. Mol. Biol.* **2012**, 422 (4), 477–487.

(28) Eisoldt, L.; Thamm, C.; Scheibel, T. Review the role of terminal domains during storage and assembly of spider silk proteins. *Biopolymers* **2012**, 97 (6), 355–361.

(29) Heidebrecht, A.; Scheibel, T. Recombinant production of spider silk proteins. *Adv. Appl. Microbiol.* **2013**, 82, 115–153.

(30) Gauthier, M.; Leclerc, J.; Lefevre, T.; Gagne, S. M.; Auger, M. Effect of pH on the Structure of the Recombinant C-Terminal Domain of Nephila clavipes Dragline Silk Protein. *Biomacromolecules* **2014**, 15 (12), 4447–4454.

(31) Huemmerich, D.; Helsen, C. W.; Quedzuweit, S.; Oschmann, J.; Rudolph, R.; Scheibel, T. Primary structure elements of spider dragline silks and their contribution to protein solubility. *Biochemistry* **2004**, 43 (42), 13604–13612.

(32) Savitzky, A.; Golay, M. J. E. Smoothing + Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, 36 (8), 1627–1739.

(33) Yang, J. T.; Wu, C. S. C.; Martinez, H. M. Calculation of Protein Conformation from Circular-Dichroism. *Methods Enzymol.* **1986**, 130, 208–269.

(34) Szabo, A.; Stolz, L.; Granzow, R. Surface-Plasmon Resonance and Its Use in Biomolecular Interaction Analysis (Bia). *Curr. Opin. Struct. Biol.* **1995**, 5 (5), 699–705.

(35) Daniels, P. B.; Deacon, J. K.; Eddowes, M. J.; Pedley, D. G. Surface-Plasmon Resonance Applied to Immunosensing. *Sens. Actuators* **1988**, 15 (1), 11–18.

(36) Klunk, W. E.; Pettegrew, J. W.; Abraham, D. J. Quantitative-Evaluation of Congo Red Binding to Amyloid-Like Proteins with a Beta-Pleated Sheet Conformation. *J. Histochem. Cytochem.* **1989**, 37 (8), 1273–1281.

(37) Semisotnov, G. V.; Rodionova, N. A.; Razgulyaev, O. I.; Uversky, V. N.; Gripas, A. F.; Gilmanshin, R. I. Study of the Molten Globule Intermediate State in Protein Folding by a Hydrophobic Fluorescent-Probe. *Biopolymers* **1991**, 31 (1), 119–128.

(38) De Simone, G.; Galdiero, S.; Manco, G.; Lang, D.; Rossi, M.; Pedone, C. A snapshot of a transition state analogue of a novel thermophilic esterase belonging to the subfamily of mammalian hormone-sensitive lipase. *J. Mol. Biol.* **2000**, 303 (5), 761–771.

(39) Manco, G.; Adinolfi, E.; Pisani, F. M.; Ottolina, G.; Carrea, G.; Rossi, M. Overexpression and properties of a new thermophilic and thermostable esterase from *Bacillus acidocaldarius* with sequence similarity to hormone-sensitive lipase subfamily. *Biochem. J.* **1998**, 332, 203–212.

(40) Vassar, P. S.; Culling, C. F. A. Fluorescent Stains, with Special Reference to Amyloid and Connective Tissues. *Arch. Pathol.* **1959**, 68 (5), 487–498.

(41) Levine, H. Thioflavine-T Interaction with Synthetic Alzheimers-Disease Beta-Amyloid Peptides - Detection of Amyloid Aggregation in Solution. *Protein Sci.* **1993**, 2 (3), 404–410.

(42) Sajid, J.; Elhaddaoui, A.; Turrell, S. Investigation of the binding of Congo red to amyloid in Alzheimer's diseased tissue. *J. Mol. Struct.* **1997**, 408–409, 181–184.

(43) Manning, M.; Colon, W. Structural basis of protein kinetic stability: resistance to sodium dodecyl sulfate suggests a central role for rigidity and a bias toward beta-sheet structure. *Biochemistry* **2004**, 43 (35), 11248–54.

(44) Jin, H. J.; Kaplan, D. L. Mechanism of silk processing in insects and spiders. *Nature* **2003**, 424 (6952), 1057–1061.

(45) Ortlepp, C. S.; Gosline, J. M. Consequences of forced silking. *Biomacromolecules* **2004**, 5 (3), 727–731.

# Supporting Information

Conformational stability and interplay of helical N-  
and C-terminal domains with implications on major  
ampullate spidroin assembly

*Joschka Bauer and Thomas Scheibel\**

Lehrstuhl Biomaterialien, Universität Bayreuth, Universitätsstr. 30, 95440 Bayreuth, Germany



**Amino acid sequences:**

**NRN1**

10 20 30 40 50  
GQANTPWSSK ANADAFINSF ISAASNTGSF SQDQMEDMSL IGNTLMAAMD  
60 70 80 90 100  
NMGGRITPSK LQALDMAFAS SVAEIAASEG GDLGVTTNAI ADALTSAFYQ  
110 120 130  
TTGVVNSRFI SEIRSLIGMF AQASANDVYA SAGSG

**NRC1**

10 20 30 40 50  
GSGPGQIYYG PQSVAAPAAA AASALAAPAT SARISSHASA LLSNGPTNPA  
60 70 80 90 100  
SISNVISNAV SQISSNPGA SACDVLVQAL LELVTALLTI IGSSNIGSVN  
110 120  
YDSSGQYAQV VTQSVQNAFA G

**Module A**

10 20  
GPYGPASAA AAAAGGYGPG SGQQ

**Module Q**

10 20  
GPGQQGPGQQ GPGQQGPGQQ

**NR3**

10 20 30 40 50  
GAASAAVSVG GYGPQSSSAP VASAAASRLS SPAASSRVSS AVSSLVSSGP  
60 70 80 90 100  
TNQAALSNTI SSVVSQVSAS NPGLSGCDVL VQALLEVVSA LVSILGSSSI  
110 120  
GQINYGASAQ YTQMVGQSVA QALAG

### **Extended experimental section:**

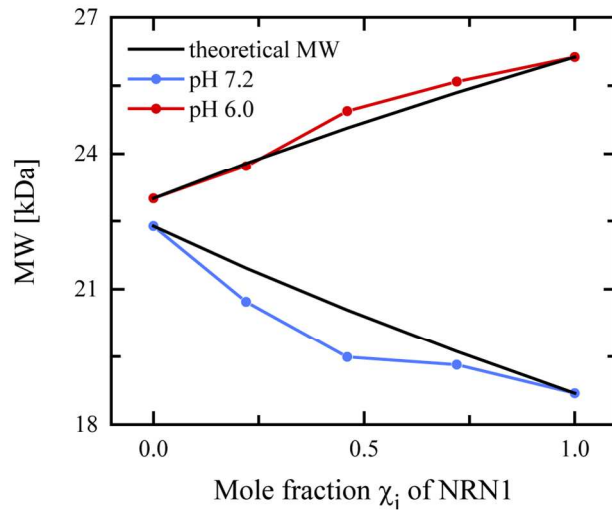
**Pull down assay.** For pull down experiments, HisPur™ Ni-NTA Spin Columns (Thermo Fisher) were washed with 20 mM NaP buffer (pH 7.2) containing 5 mM imidazole while gently shaking at 350 rpm at RT. The columns were activated with 0.5 mM NiSO<sub>4</sub> and washed. His-SUMO tagged NRN1 (51.0 μM) was bound to the column upon incubation for 1 h at RT. After unbound proteins were washed off, binding between NRC1 (267.8 μM) and immobilized SUMO-NRN1 was tested at distinct pH values (pH 7.2 or pH 6.0) and salt conditions (w/o salt, 200 mM NaCl, 200 mM KH<sub>2</sub>PO<sub>4</sub>/K<sub>2</sub>HPO<sub>4</sub>). Nonspecifically bound proteins were removed upon four washing steps with appropriate buffers. Bound proteins were eluted with 50 mM Tris (pH 7.5) containing 300 mM imidazole as well as 100 mM NaCl and separated by SDS-PAGE using a pre-cast 4 - 20 % acrylamide gradient gel (SERVAGel™ TG PRiME™) followed by Coomassie blue staining.

**Native PAGE.** NRN1 and NRC1 were mixed at distinct molar ratios (1:0, 3:1, 1:1, 1:3, 0:1) at a total protein concentration of 70 μM and incubated for 10 min at RT. Sample buffer with 25 mM BisTris (pH 7.2) and 30 % (w/v) glycerol was added, and samples (5 μg) were analyzed by native PAGE using a precast 4 - 20 % acrylamide gradient gel (SERVAGel™ TG PRiME™). For evaluation, the gel was stained using Coomassie blue.

**Size exclusion chromatography multi-angle light scattering (SEC-MALS).** Mixtures of 2 mg ml<sup>-1</sup> NRN1 (142.5 μM) and NRC1 (167.4 μM) were pre-separated by size-exclusion chromatography (SEC) (Superdex™ 200 10/300 GL, GE Healthcare). An Agilent 1100 HPLC system was employed using 20 mM sodium phosphate buffer (pH 7.2 or 6.0) at a flow rate of 0.7 ml min<sup>-1</sup>, and 250 μl of the protein mixture were injected at a flow rate of 0.2 ml min<sup>-1</sup>. The molecular weight (MW) was subsequently analyzed by multi-angle light scattering (MALS) and

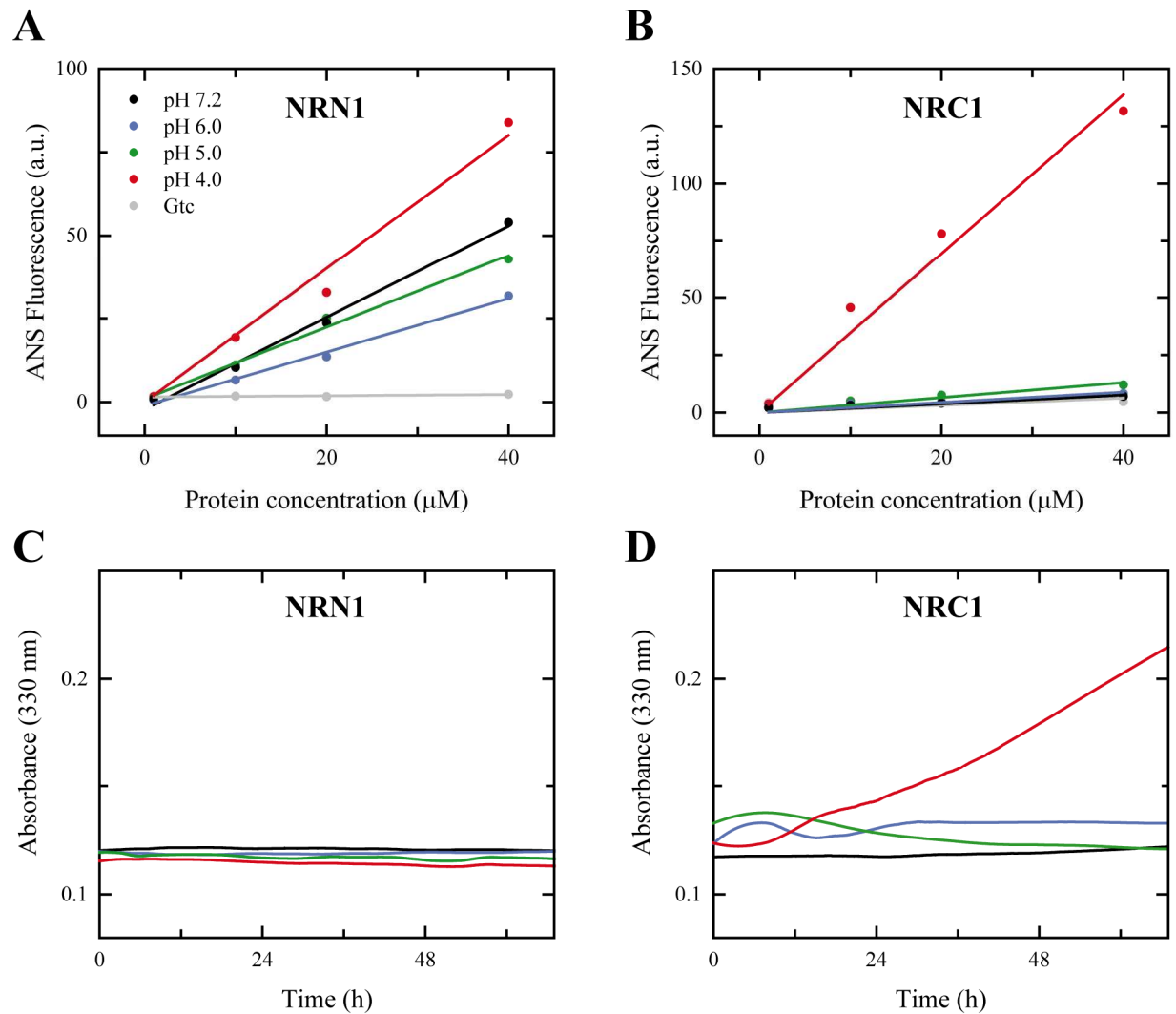
quasi-elastic light scattering (QELS, WYATT). The data were processed using the ASTRA software (WYATT). The mole fraction ( $\chi_i$ ) of NRN1 was calculated by applying  $\chi_i = \frac{n(NRN1)}{n(NRN1)+n(NRC1)}$ . Endpoint values of 0 and 1, representing 100 % NRC1 and 100 % NRN1, respectively, were used to determine the theoretical MW.

### **Additional figures:**

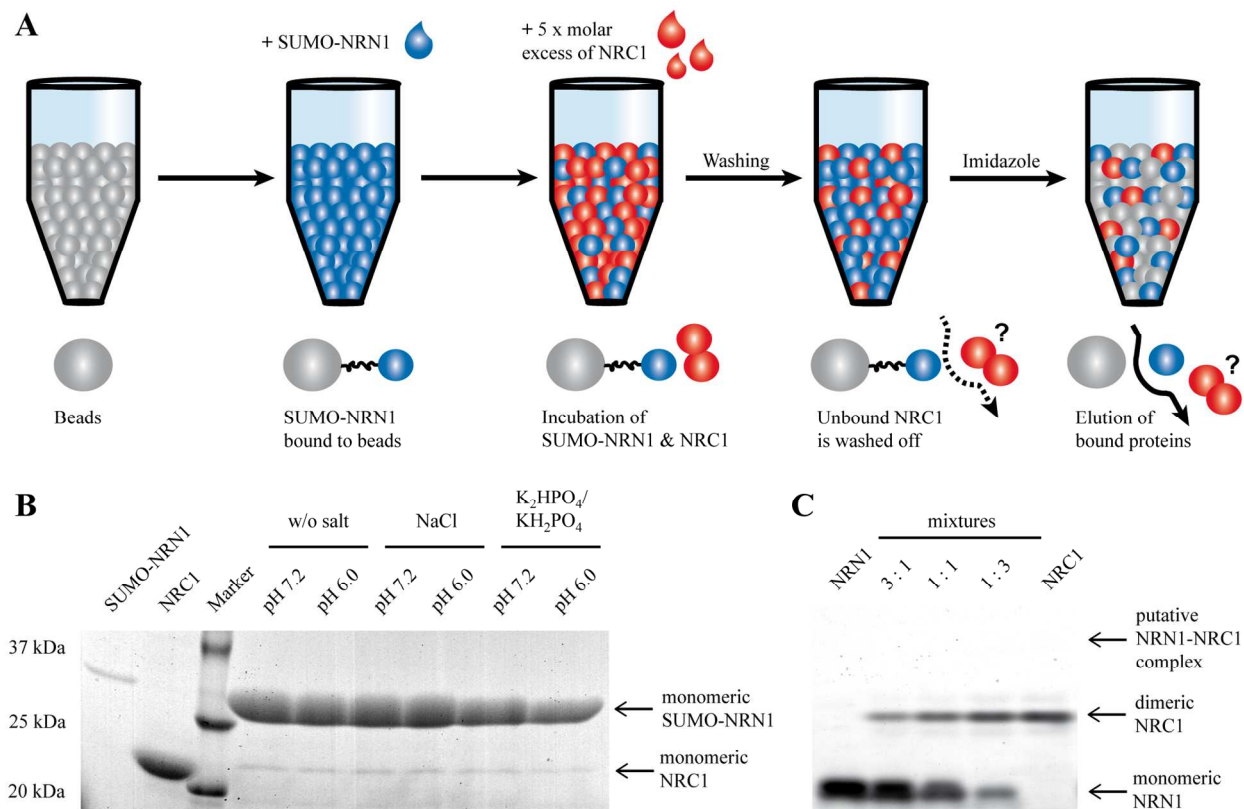


**Figure S1:** Absolute MWs of NRN1-NRC1 samples mixed at different ratios as determined by MALS. The average MW of NRN1 (monomer/dimer, 14/28 kDa) and NRC1 (dimer, 24 kDa) mixtures were analyzed at pH 7.2 and pH 6.0. Proteins were separated by SEC prior to MALS analysis. Endpoint values of  $\chi_i = 0$  and  $\chi_i = 1$ , representing 100 % (w/w) NRC1 and 100 % (w/w) NRN1, respectively, were used for defining the theoretical MW of mixtures. The experimentally determined MW of NRN1-NRC1 mixtures significantly correlated with the theoretical MW of a non-interacting protein mixture, indicating that the terminal spidroin domains did apparently not interact.

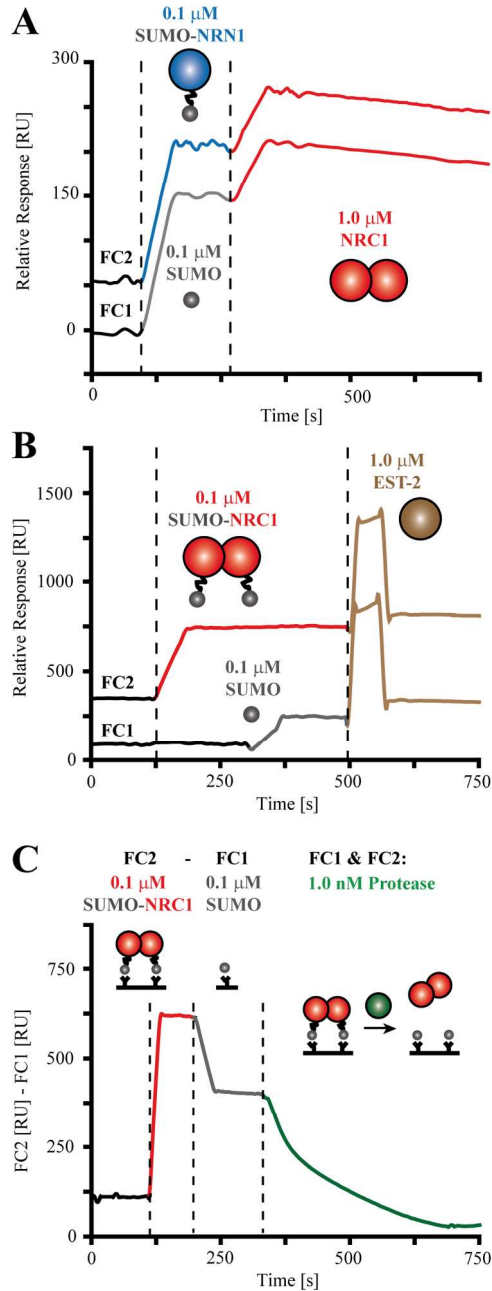




**Figure S2:** Aggregation analysis of NRN1 and NRC1. The color coding scheme is shown in part A of the figure. **A, B)** Dependence of ANS fluorescence at 475 nm on **(A)** NRN1 and **(B)** NRC1 concentration between pH 7.0 and 4.0. Linear fits are marked by solid lines. Emission of ANS in the presence of denatured protein (6 M guanidinium thiocyanate, Gtc) served as negative control. **C, D)** Turbidity of  $0.5 \text{ mg ml}^{-1}$  **(C)** NRN1 and **(D)** NRC1 was recorded at a wavelength of 330 nm in a time-dependent manner ( $n = 3$ ).

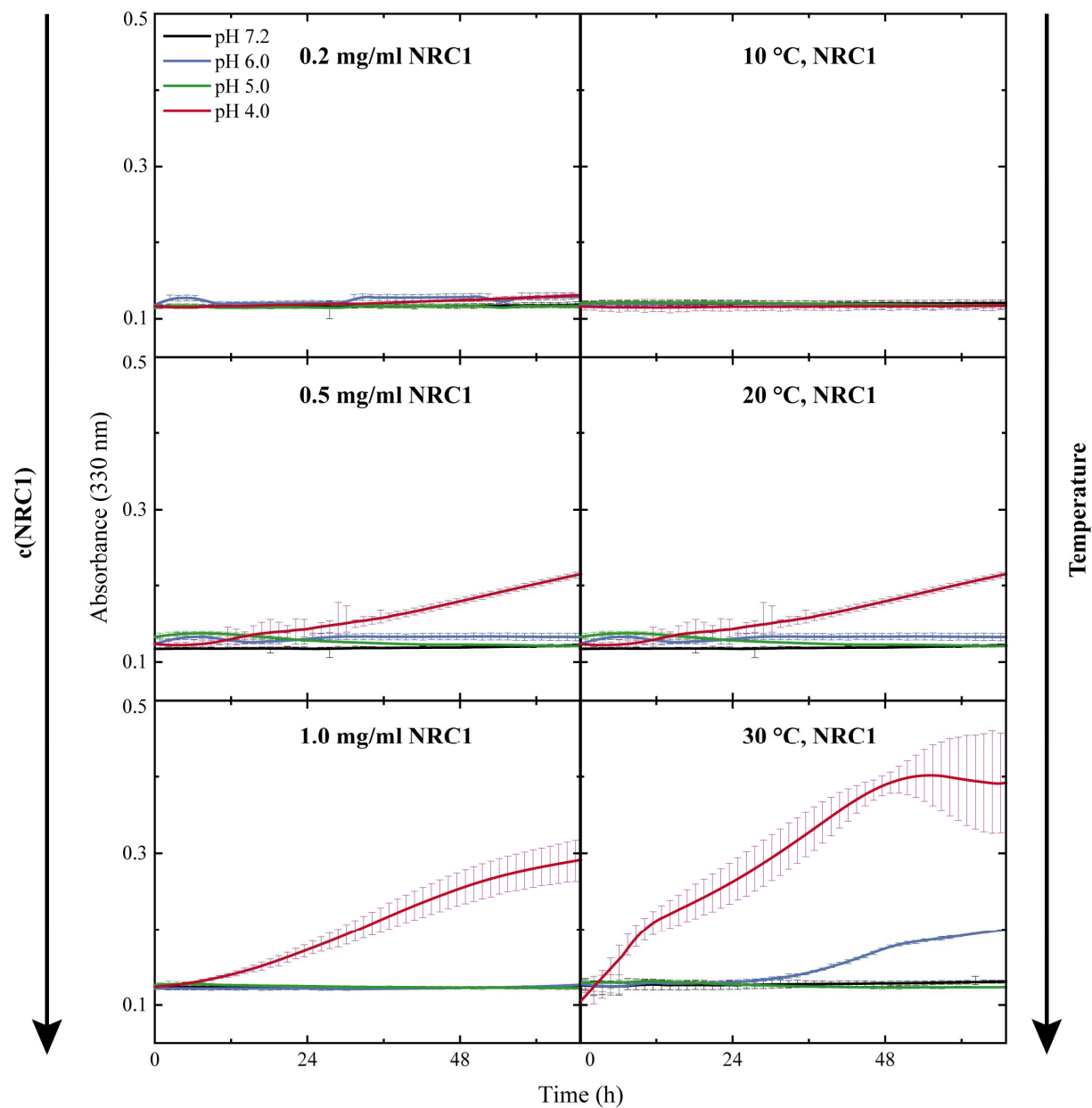


**Figure S3:** Interaction analysis between monomeric NRN1 and dimeric NRC1. **A)** Scheme of the pull-down assay. After binding of 51.0  $\mu$ M monomeric SUMO-NRN1 to Ni-NTA beads, a fivefold molar excess of dimeric NRC1 was added at distinct solvent conditions (pH 7.2/6.0  $\pm$  200 mM NaCl or  $\pm$   $K_2HPO_4/KH_2PO_4$ ), and the mixture was incubated for 1 h. The protein-loaded beads were washed four times, and bound proteins were eluted in the presence of imidazole. **B)** The eluted protein mix was separated using a reducing SDS-PAGE (i.e. if NRC1 is bound to the beads it will run as a monomer in this gel). The 4-20 % acrylamide gradient gel was subsequently stained with Coomassie blue. **C)** Native PAGE was performed to test NRN1-NRC1 interactions at distinct molar ratios (3:1, 1:1 and 1:3). While NRN1 was monomeric at the tested pH 7.2, NRC1 formed disulfide-linked dimers. However, no NRN1-NRC1 complexes could be detected at these conditions. The gel was subsequently stained with Coomassie blue.

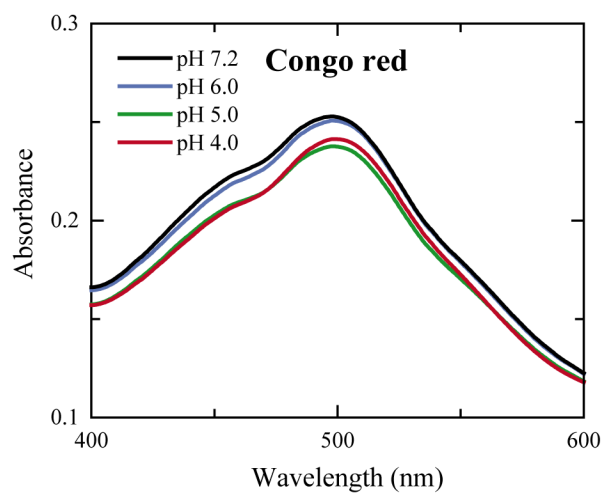


**Figure S4:** Surface plasmon resonance experiments. **A)** SPR analysis of the interaction between 0.1  $\mu$ M SUMO-NRN1 monomers (FC2) and 1.0  $\mu$ M dimeric NRC1. FC1 was used as negative control with immobilized 0.1  $\mu$ M SUMO. **B)** Binding of 1.0  $\mu$ M esterase-2 to 0.1  $\mu$ M SUMO-NRN1 (FC2) and to 0.1  $\mu$ M SUMO (FC1) was measured as a negative control. **C)** NRC1 was cleaved off the SUMO-NRC1 construct using the protease ULPI (1 nM). The response of FC1 (0.1  $\mu$ M SUMO) is subtracted from the response of FC2 (0.1  $\mu$ M SUMO-NRC1).

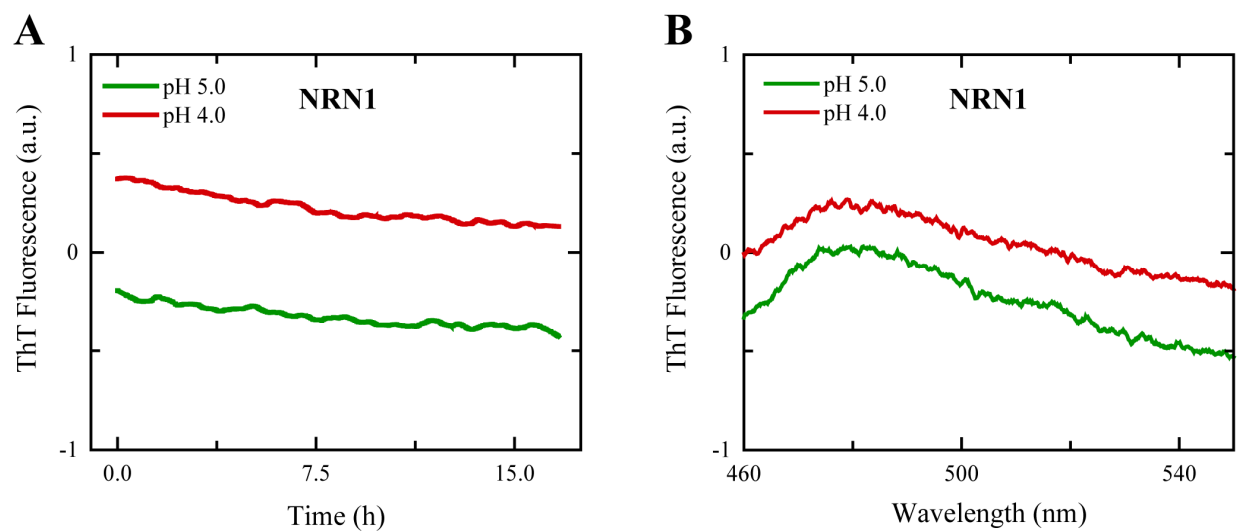




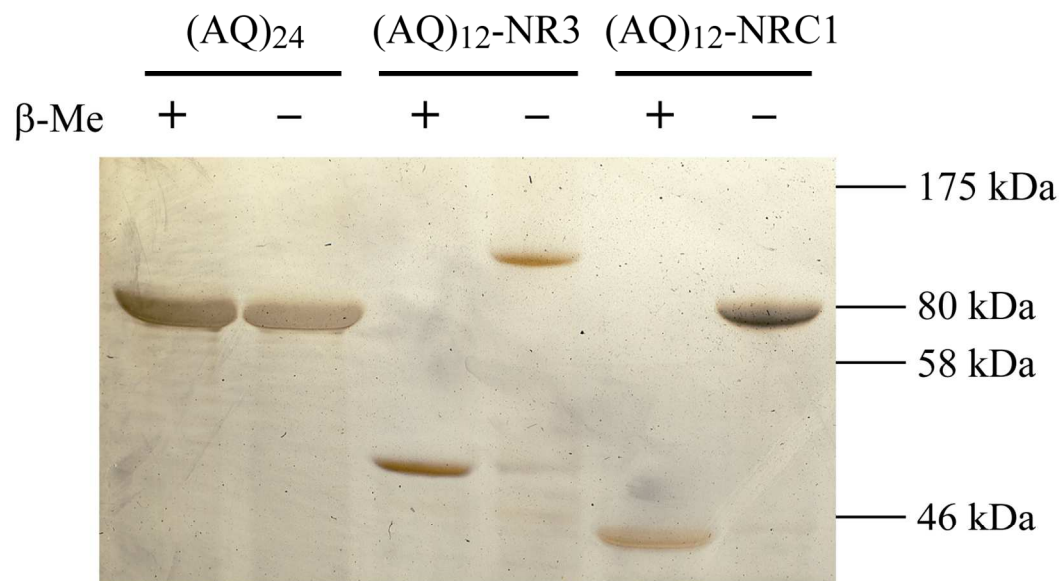
**Figure S5:** Turbidity of NRC1 in dependence of protein concentration and temperature. The absorbance was recorded at a wavelength of 330 nm ( $n = 3$ ). Every fiftieth error bar is shown, and a Savitzky-Golay filter is applied.<sup>1</sup>



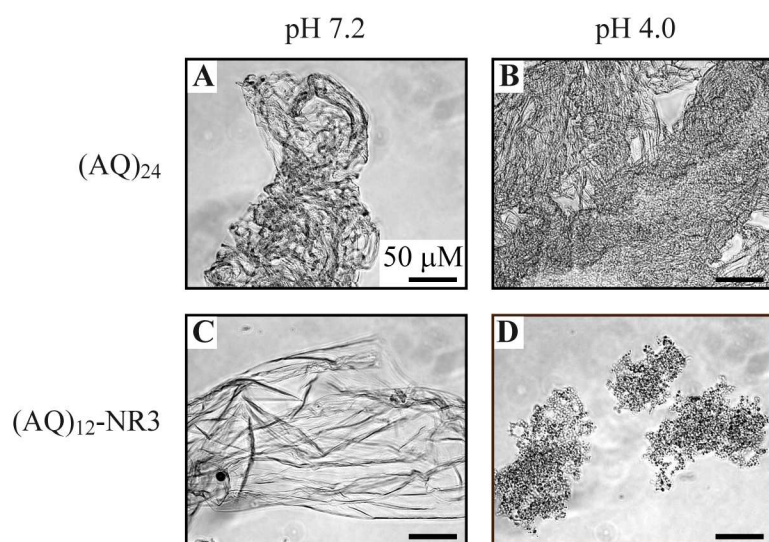
**Figure S6:** Absorbance of 4  $\mu$ M Congo red recorded at distinct pH values.



**Figure S7: A)** Time-dependent fluorescence changes of 5  $\mu$ M Thioflavine T were measured at a wavelength of 482 nm in the presence of 40  $\mu$ M NRN1. **B)** Fluorescence spectra were taken after incubating ThT and NRN1 for 17 h.



**Figure S8:** Dimerization analysis of recombinant spidroins using SDS PAGE. β-mercaptoethanol reduces the intermolecular disulfide-bridge in the carboxy-terminal domains yielding monomers in case of (AQ)<sub>12</sub>-NR3 and (AQ)<sub>12</sub>-NRC1. Although both constructs exhibit a similar MW (Table S1), the propagation in the PAGE differed due to differences in the amino acid composition. The 4-12 % gradient gel was silver stained after electrophoresis.



**Figure S9:** Bright-field microscopy images of  $(AQ)_{24}$  (A, B) and  $(AQ)_{12}$ -NRC1 aggregates (C, D). The pH was adjusted to 7.2 (A, C) or 4.0 (B, D), and the samples were rotated for 18 h at 25 r.p.m.

#### Additional tables:

**Table S1:** The percentage of secondary structure elements was estimated for NRC1 by using JASCO Spectra Manager<sup>TM</sup> II and matching the CD spectra to Yang's references.<sup>2</sup>

	$\alpha$ -helices	$\beta$ -sheets	random coils
<b>pH 7.2</b>	56.3	18.5	23.4
<b>pH 6.0</b>	43.9	27.7	27.8
<b>pH 5.0</b>	41.8	30.8	27.4
<b>pH 4.0</b>	16.5	47.1	34.0

Table S2: The percentage of secondary structure elements was estimated for mixtures of NRN1 and NRC1 by using JASCO Spectra Manager<sup>TM</sup> II and matching the CD spectra to Yang's references.<sup>2</sup>

	$\alpha$ -helices	$\beta$ -sheets	random coils
<b>pH 7.2</b>	51.6	19.9	28.4
<b>pH 6.0</b>	52.3	23.4	23.9
<b>pH 5.0</b>	25.3	38.2	20.4
<b>pH 4.0</b>	26.3	43.4	26.3

**Table S3:** Physico-chemical properties of recombinant spidroins.

	NRC1	(AQ) <sub>24</sub>	(AQ) <sub>12</sub> -NR3	(AQ) <sub>12</sub> -NRC1
<b>MW (monomers)</b>	12.0 kDa	94.5 kDa	58.4 kDa	59.7 kDa
<b>theoretical pI</b>	4.5	7.6	5.8	5.4
<b>number of negatively charged residues</b>	3	0	2	3
<b>number of positively charged residues</b>	1	1	2	2

## REFERENCES

1. Savitzky, A.; Golay, M. J. E., Smoothing + Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964**, 36, (8), 1627-1739.
2. Yang, J. T.; Wu, C. S. C.; Martinez, H. M., Calculation of Protein Conformation from Circular-Dichroism. *Method. Enzymol.* **1986**, 130, 208-269.

**DANKSAGUNG**

Als erstes möchte ich mich herzlich bei meinem Mentor Herrn Prof. Dr. Thomas Scheibel für die Bereitstellung der interessanten Themen und die exzellente Betreuung meiner Arbeit bedanken. Durch die kameradschaftliche Unterstützung konnte ich in den verschiedensten Bereichen von ihm lernen, gleichzeitig wurden mir genügend Freiheiten gelassen, um auch eigene Projektideen umzusetzen. Insbesondere möchte ich mich für die großzügigen Möglichkeiten bedanken, durch die Leitung der Lehrstuhl-eigenen IT Administration und des Kooperationsprojekts mit der Messersmith Group (University of California) Verantwortung zu übernehmen.

Weiterhin möchte ich mich bei Herrn Prof. Dr. Stephan Schwarzingen und Herrn Prof. Dr. Franz Xaver Schmid für die Übernahme meines Mentorats bedanken.

Ein großer Dank für die erfolgreiche Kooperation gilt insbesondere Daniel Schaal, Kristian Schweimer und Stephan Schwarzingen von dem Lehrstuhl Biopolymere, in deren Rahmen die Struktur der Variante NRN1-3\*<sub>L.h.</sub> von *Latrodectus hesperus* MaSp1 gelöst wurde.

Folgenden Kollegen vom Lehrstuhl Biomaterialien möchte ich für deren Unterstützung danken:

- Andreas Schmidt und Johannes Diehl für die Hilfe bei der Fermentation
- Christian Borkner für REM-Aufnahmen und die Unterstützung bei SPR- und FTIR-Spektroskopie
- Kristin Schacht für TEM-Aufnahmen
- Elene Doblhofer für die Unterstützung bei FTIR Messungen und fotografischen Aufnahmen
- Tamara Aigner für die Hilfe bei TEM und Durchlichtmikroskopie
- Dr. Hendrik Bargel für die REM-Aufnahmen
- Dr. Martin Humenik für AFM-Aufnahmen und Messungen am Massenspektrometer
- Meinen Praktikanten Corinna Drescher, Svenja Schiermann, Bernd Steiger und Fabian Müller
- Elise DeSimone und Dr. Martin Humenik für das Korrekturlesen von Manuskripten
- Andreas Schmidt, Johannes Diehl, Alexandra Pellert, Claudia Stemmann und Eva Möller für die Unterstützung des Laboralltags
- Meiner Arbeitsgruppe „Proteinfaltung“ für die hilfreichen Diskussionen



- Meinen Bürokollegen Christopher Thamm, Christian Haynl, Jia Wang, Christian Borkner, Fabian Müller und Michael Suhre für die vielen wissenschaftlichen und privaten Gespräche und der freundschaftlichen Atmosphäre im Büro
- Susanne Schramm, Andrea Bodner und Sabrina Raum für die Hilfe bei organisatorischen Angelegenheiten
- Allen aktuellen und ehemaligen Kollegen des gesamten Lehrstuhls Biomaterialien für die tolle Arbeitsatmosphäre

Ein besonderer Dank gilt Dr. Martin Humenik für die zahlreichen wissenschaftlichen Diskussionen, die einen wertvollen Beitrag zu den Veröffentlichungen und zu dieser Arbeit geleistet haben. Zudem danke ich Dr. Martin Humenik, Christopher Thamm und Martin Pannek für das Korrekturlesen dieser Arbeit.

Meinen Freunden Christopher Thamm, Christian Haynl, Andreas Schmidt, Fabian Müller, Elena Doblhofer, Kristin Schacht, Martin Pannek, Manuel Mönnich und Manuel Braun ein großes Dankschön für die schöne Zeit in Bayreuth und die interessanten Gespräche beim Mittagessen.

Zuletzt möchte ich mich herzlichst bei meiner Familie für die endlose Unterstützung bedanken. Meine Eltern Volker und Dagmar haben mir in jeder Situation Kraft gespendet und Vertrauen geschenkt. Die gemeinsame Zeit mit meinen Geschwistern Viola und Henrik bot mir immer eine willkommene Abwechslung. Ein ganz spezieller Dank geht an meine Frau Tamara! Vielen lieben Dank dafür, dass du mir trotz den vielen zeitlichen Engpässen so selbstlos den Rücken freigehalten, mich hingebungsvoll unterstützt und so viel Freude und Leichtigkeit in mein Leben gebracht hast.

**EIDESSTATTLICHE VERSICHERUNGEN UND ERKLÄRUNGEN**

(§ 9 Satz 2 Nr. 3 PromO BayNAT)

Hiermit versichere ich eidesstattlich, dass ich die Arbeit selbständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe (vgl. Art. 64 Abs. 1 Satz 6 BayHSchG).

(§ 9 Satz 2 Nr. 3 PromO BayNAT)

Hiermit erkläre ich, dass ich die Dissertation nicht bereits zur Erlangung eines akademischen Grades eingereicht habe und dass ich nicht bereits diese oder eine gleichartige Doktorprüfung endgültig nicht bestanden habe.

(§ 9 Satz 2 Nr. 4 PromO BayNAT)

Hiermit erkläre ich, dass ich Hilfe von gewerblichen Promotionsberatern bzw. -vermittlern oder ähnlichen Dienstleistern weder bisher in Anspruch genommen habe noch künftig in Anspruch nehmen werde.

(§ 9 Satz 2 Nr. 7 PromO BayNAT)

Hiermit erkläre ich mein Einverständnis, dass die elektronische Fassung meiner Dissertation unter Wahrung meiner Urheberrechte und des Datenschutzes einer gesonderten Überprüfung unterzogen werden kann.

(§ 9 Satz 2 Nr. 8 PromO BayNAT)

Hiermit erkläre ich mein Einverständnis, dass bei Verdacht wissenschaftlichen Fehlverhaltens Ermittlungen durch universitätsinterne Organe der wissenschaftlichen Selbstkontrolle stattfinden können.

Bayreuth, 24.06.2018, .....